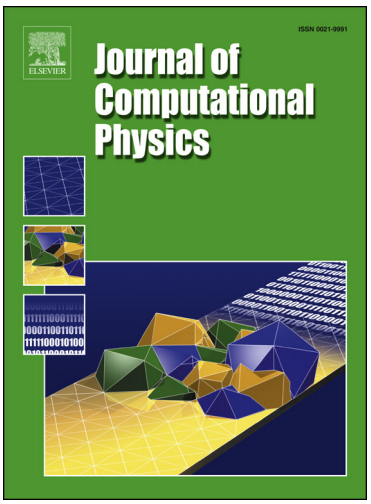
Accepted Manuscript



Gamblets for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic ODEs/PDEs with rough coefficients

Houman Owhadi, Lei Zhang

PII: S0021-9991(17)30491-6
DOI: http://dx.doi.org/10.1016/j.jcp.2017.06.037
Reference: YJCPH 7436

To appear in: *Journal of Computational Physics*

Received date: 24 June 2016
Revised date: 16 June 2017
Accepted date: 24 June 2017

Please cite this article in press as: H. Owhadi, L. Zhang, Gamblets for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic ODEs/PDEs with rough coefficients, *J. Comput. Phys.* (2017), http://dx.doi.org/10.1016/j.jcp.2017.06.037

Gamblets for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic ODEs/PDEs with rough coefficients

Houman Owhadi*and Lei Zhang†

June 27, 2017

Abstract

Implicit schemes are popular methods for the integration of time dependent PDEs such as hyperbolic and parabolic PDEs. However the necessity to solve corresponding linear systems at each time step constitutes a complexity bottleneck in their application to PDEs with rough coefficients. We present a generalization of gamblets introduced in [62] enabling the resolution of these implicit systems in near-linear complexity and provide rigorous a-priori error bounds on the resulting numerical approximations of hyperbolic and parabolic PDEs. These generalized gamblets induce a multiresolution decomposition of the solution space that is adapted to both the underlying (hyperbolic and parabolic) PDE (and the system of ODEs resulting from space discretization) and to the time-steps of the numerical scheme.

1 Introduction

Implicit schemes are popular and powerful methods for the integration of time dependent PDEs such as hyperbolic and parabolic PDEs [96, 44, 43, 11]. However the necessity to solve corresponding linear systems at each time step constitutes a complexity bottleneck in their application to PDEs with rough coefficients.

*California Institute of Technology, Computing & Mathematical Sciences , MC 9-94 Pasadena, CA 91125, owhadi@caltech.edu

†Shanghai Jiao Tong University, School of Mathematical Sciences, Institute of Natural Sciences, and Ministry of Education Key Laboratory of Scientific and Engineering Computing (MOE-LSC), Shanghai, 200240, China, lzhang2012@sjtu.edu.cn

Although multigrid methods [34, 12, 36] have been successfully generalized to time dependent equations [50, 97, 96, 110, 33, 106, 44], their convergence rate can be severely affected by the lack of regularity of the coefficients [102]. While some degree of robustness can be achieved with algebraic multigrid [80], multilevel finite element splitting [112], hierarchical basis multigrid [6], multilevel preconditioning [98], stabilized hierarchical basis methods [99] and energy minimization [52, 102, 109], the design of multigrid/multiresolution methods that are provably robust with respect to rough (L^∞) coefficients was an open problem of practical importance [13] addressed in [62] with the introduction of gamblets (in $\mathcal{O}(N\ln^{3d}N)$ complexity for the first solve and $\mathcal{O}(N\ln^{d+1}N)$ for subsequent solves to achieve grid-size accuracy in H^1-norm for elliptic problems). Numerical evidence suggests the robustness of low rank matrix decomposition based methods such as the Fast Multipole Method [35, 111], Hierarchical matrices [37, 7] and Hierarchical Interpolative Factorization [42] and while this robustness can be proven rigorously for Hierarchical matrices [7] (the complexity of Hierarchical matrices is $\mathcal{O}(N\ln^{2d+8}N)$ to achieve grid-size accuracy in L^2-norm for elliptic problems [7]) one may wonder if it is possible to rigorously lower this known complexity bound and achieve (at the same time) a meaningful multi-resolution decomposition of the solution space for time dependent problems. Although classical wavelet based methods [14, 10, 28] enable a multi-resolution decomposition of the solution space their performance is also affected by the regularity of coefficients because they are not adapted to the underlying PDEs.

In section 2 we present a generalization of gamblets introduced in [62] and apply them in sections 3 and 4 to the implicit schemes for hyperbolic and parabolic PDEs with rough coefficients. As in [62] these generalized gamblets (1) are elementary solutions of hierarchical information games associated with the process of computing with partial information and limited resources, (2) have a natural Bayesian interpretation under the mixed strategy emerging from the game theoretic formulation, (3) induce a multi-resolution decomposition of the solution space that is adapted to the space-time numerical discretization of the underlying PDE and propagate the solution independently (at each time-step) in each sub-band of the decomposition. The complexity of pre-computing generalized gamblets is $N\ln^{3d}N$ and that of propagating the solution is $N\ln^{d+1}N$ (at each time step, to achieve grid-size accuracy in energy norm). Although real valued gamblets are sufficient for first and second order implicit schemes, higher order implicit schemes may require complex valued gamblets. These complex valued gamblets are introduced and their application to higher order schemes is illustrated in

Section 5. Observing that the multiresolution decomposition induced by gamblets has properties that are similar to an eigenspace decomposition, we introduce, in Section 6, a multi-time-step scheme for solving parabolic PDEs (with rough coefficients) in $\mathcal{O}(N \ln^{3d+1} N)$ complexity.

Gamblets are derived from a Game Theoretic approach to Numerical Analysis [62, 66] which could be seen as decision theory approach to numerical analysis [101, 65]. We refer to the information based complexity literature for an understanding of the natural connection between the notions of computing with partial/priced information and numerical analysis (we refer in particular to [107, 74, 95, 55, 108, 79, 56]). Although statistical approaches to numerical analysis [26, 77, 89, 48, 81, 47, 85, 57, 58] have, in the past, received little attention, perhaps due to the counterintuitive nature of the process of randomizing a *known* function, the possibilities offered by combining numerical uncertainties/errors with model uncertainties/errors appear to be stimulating their reemergence [19, 84, 61, 41, 40, 15, 23, 78, 75, 66, 82]. We refer in particular to [86, 84, 19] for ODEs and to [61, 62, 66, 20, 83] for PDEs. Here the game theoretic approach of [62] is applied to both PDEs and the system of ODEs resulting from their discretization. The multiscale nature of the underlying PDEs results in the stiffness of the corresponding ODEs (these ODEs are non only stiff [92, 94, 93] they are also characterized by a large range/continuum of time scales [59, 60, 8]). Although it is natural to integrate such ODEs by an eigenspace decomposition when the dimension of the system of ODEs is small, the cost of such an approach is in general prohibitive. It is to some degree surprising that gamblets have properties that are similar to eigenfunctions, or more precisely Wannier basis functions [104, 53] (i.e. linear combinations of eigenfunctions concentrated around a given eigenvalue that are also concentrated in space), while preserving the near-linear complexity of the integration.

Since (see [62]) Gamblets are also natural basis functions for numerical homogenization [105, 3, 46, 30, 70, 31, 9, 2, 25, 91, 103, 72, 51, 73, 45, 76] they can also be employed to achieve sub-linear complexity under sufficient regularity of source terms and initial conditions (see [71, 69, 72] and Remark 4.3).

We also refer to [66] for a generalization of gamblets to arbitrary continuous linear bijections on Banach spaces (see also [83] for their application to the inversion, compression and approximate PCA of dense kernel matrices at near-linear complexity). As discussed in [66] gamblets also provide a solution to the problem of identifying operator adapted wavelets [21, 4, 32, 17, 18, 24, 1, 90, 100, 88] satisfying three essential properties

(see [87, 88] for an overview): (a) scale-orthogonality (with respect to the operator scalar product to ensure block-diagonal stiffness matrices) (b) local support (or rapid decay) of the wavelets (to ensures that the individual blocks are sparse) and (c) Riesz stability in the energy norm (to ensure that the blocks are well-conditioned).

2 Gamblets

We will, in this section, present a generalization of the gamblets introduced in [62]. Since the proofs of the results presented in this section are similar to those given in [62] we will refer the reader to [62] and to [66] for these proofs.

2.1 The PDE

Let $\zeta > 0$. Consider the PDE

$$\begin{cases} \frac{4}{\zeta^2}\mu(x)u(x) - \operatorname{div}\big(a(x)\nabla u(x)\big) = g(x) \quad x \in \Omega; \\ u = 0 \quad \text{on} \quad \partial\Omega, \end{cases} \tag{2.1}$$

where Ω is a bounded domain in $\mathbb{R}^d$ (of arbitrary dimension $d \in \mathbb{N}^*$) with piecewise Lipschitz boundary, a is a symmetric, uniformly elliptic $d \times d$ matrix with entries in $L^\infty(\Omega)$ and such that for all $x \in \Omega$ and $l \in \mathbb{R}^d$,

$$\lambda_{\min}(a)|l|^2 \leq l^T a(x) l \leq \lambda_{\max}(a)|l|^2, \tag{2.2}$$

and $\mu \in L^\infty(\Omega)$ with for all $x \in \Omega$,

$$\mu_{\min} \leq \mu(x) \leq \mu_{\max}\,. \tag{2.3}$$

One purpose of gamblets is to compute the solution of (2.1) (or its finite-element solution) as fast as possible to a given accuracy.

2.2 The hierarchy of measurement functions

We will now introduce a hierarchy of measurement functions that will be used to characterize the process of computing of hierarchies of levels of complexity. We will need the following hierarchy of labels.

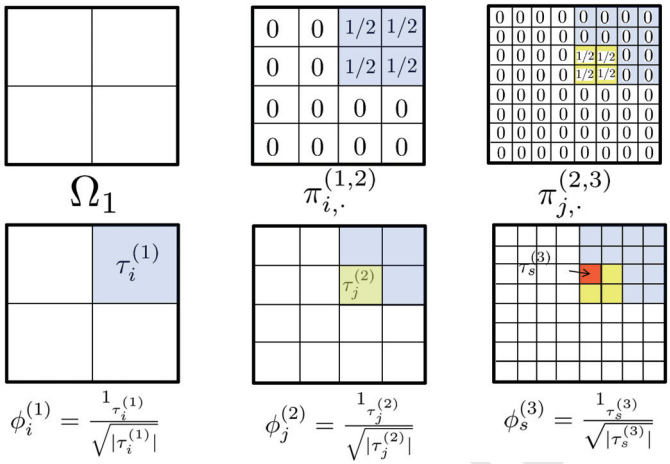


Figure 1: $\Omega = (0,1)^2$. Ω_k corresponds to a uniform partition of Ω into $2^{-k} \times 2^{-k}$ squares. The bottom row shows the support of $\phi_i^{(1)}, \phi_j^{(2)}$ and $\phi_s^{(3)}$. Note that $j^{(1)} = s^{(1)} = i$ and $s^{(2)} = j$. The top row shows the entries of $\pi_{i,\cdot}^{(1,2)}$ and $\pi_{j,\cdot}^{(2,3)}$.

Definition 2.1. *We say that $\mathcal{I}^{(r)}$ is an index tree of depth r if it is the finite set of r-tuples of the form $i = (i_1, \ldots, i_r)$. For $1 \leq k \leq r$ and $i = (i_1, \ldots, i_r) \in \mathcal{I}^{(r)}$, write $i^{(k)} := (i_1, \ldots, i_k)$ and $\mathcal{I}^{(k)} := \{i^{(k)} : i \in \mathcal{I}^{(r)}\}$. For $1 < s < k$ and a k-tuple of the form $i = (i_1, \ldots, i_k)$ we write $i^{(s)} := (i_1, \ldots, i_s)$.*

Write $I^{(k)}$ for the $\mathcal{I}^{(k)} \times \mathcal{I}^{(k)}$ identity matrix.

Construction 2.2. *For $k \in \{1, \ldots, q-1\}$ let $\pi^{(k,k+1)}$ be a $\mathcal{I}^{(k)} \times \mathcal{I}^{(k+1)}$ matrix such that $\pi^{(k,k+1)}(\pi^{(k,k+1)})^T = I^{(k)}$ and $\pi_{i,j}^{(k,k+1)} = 0$ for $j^{(k)} \neq i$ (we say that $\pi^{(k,k+1)}$ is cellular).*

Let $(\phi_i^{(r)})_{i \in \mathcal{I}^{(r)}}$ be orthonormal elements of $L^2(\Omega)$ and, for $k \in \{1, \ldots, r-1\}$ and $i \in \mathcal{I}^{(k)}$ define $\phi_i^{(k)}$ via induction by

$$\phi_i^{(k)} = \sum_{j \in \mathcal{I}^{(k+1)}} \pi_j^{(k,k+1)} \phi_j^{(k+1)} \tag{2.4}$$

We will refer to the elements $\phi_i^{(k)}$ as measurement functions. Through this paper we use Haar wavelets or approximations thereof (Construction

2.3) as prototypical measurement functions. We refer the reader to [66] for a comprehensive description of the framework.

Construction 2.3. *Let $H, \delta \in (0,1)$. For $k \in \mathcal{N}^*$, let Ω_k be a nested partition of Ω into subsets $(\tau_i^{(k)})_{i \in \mathcal{I}^{(k)}}$ such that (1) each $\tau_i^{(k)}$ is contained in a ball of radius H and contains a ball of radius δH and (2) $|\tau_i^{(k)}| = |\tau_j^{(k)}|$ ($|\tau_i^{(k)}|$ is the volume of $\tau_i^{(k)}$). Let $\phi_i^{(k)} = \frac{1_{\tau_i^{(k)}}}{\sqrt{|\tau_i^{(k)}|}}$ where $1_{\tau_i^{(k)}}$ is the indicator function of $\tau_i^{(k)}$. Observe that the nesting matrices $\pi^{(k,k+1)}$ are cellular and orthonormal (in the sense that $\pi^{(k,k+1)}(\pi^{(k,k+1)})^T = I^{(k)}$ where $I^{(k)}$ is the $\mathcal{I}^{(k)} \times \mathcal{I}^{(k)}$ identity matrix).*

Example 2.1. *For our running example we will consider $\Omega = (0,1)^2$ illustrated in Figure 1 (taken from [66]). Using the Construction 2.3 we select Ω_k to be a regular grid partition of Ω into $2^{-k} \times 2^{-k}$ squares $\tau_i^{(k)}$ and $\phi_i^{(k)} = \frac{1_{\tau_i^{(k)}}}{\sqrt{|\tau_i^{(k)}|}}$.*

2.3 The hierarchy of games

Gamblets are then identified by turning the process of computing with limited resources and partial information as that of playing hierarchies of games defined as follows. We have two players I and II. Player I chooses the right hand side g of (2.1) in $H^{-1}(\Omega)$ and does not show it to Player II. Starting with $k = 1$, Player II sees $(\int_\Omega u\phi_i^{(k)})_{i \in \mathcal{I}^{(k)}}$ and must predict u and $(\int_\Omega u\phi_i^{(k+1)})_{i \in \mathcal{I}^{(k+1)}}$. Once Player II has made his choice, he gets a loss, sees $(\int_\Omega u\phi_i^{(k+1)})_{i \in \mathcal{I}^{(k+1)}}$ and must predict u and $(\int_\Omega u\phi_i^{(k+2)})_{i \in \mathcal{I}^{(k+2)}}$. In this adversarial game Player I tries to maximize the loss of Player II and Player II tries to minimize it. Optimal strategies are identified by lifting this deterministic minimax problem to a minimax over measures [66, Sec. 5]. In other words, Player I must play at random and Player II must look for an optimal strategy in the Bayesian class of strategies by considering the SPDE

$$\begin{cases} \frac{4}{\zeta^2}\mu v - \operatorname{div}\left(a\nabla v\right) = \xi \quad x \in \Omega; \\ u = 0 \quad \text{on} \quad \partial\Omega, \end{cases} \tag{2.5}$$

where the right hand side of (2.1) has been replaced by a random field ξ and the bet of Player II at step k is the expectation of the solution of the SPDE (2.5) conditioned on measurements of the solution of the deterministic PDE

(2.1), i.e.

$$u^{(k),\zeta}(x) := \mathbb{E}\big[v(x)\big| \int_\Omega v(y)\phi_i^{(k)}(y)\,dy = \int_\Omega u(y)\phi_i^{(k)}(y)\,dy,\ i \in \mathcal{I}^{(k)}\big]\,. \quad (2.6)$$

Note that the sequence of approximations (2.6) is a martingale under the filtration formed by the measurements $(\int_\Omega u\phi_i^{(k)})_{i\in\mathcal{I}^{(k)}}$.

2.4 ζ-Gamblets

If the loss of Player II is measured using relative error in the energy norm $\|w\|_\zeta^2 := \frac{4}{\zeta^2}\int_\Omega w^2\mu + \int_\Omega (\nabla w)^T a \nabla w$ associated with the operator scalar product

$$\big\langle w_1, w_2\big\rangle_\zeta := \frac{4}{\zeta^2}\int_\Omega w_1 w_2 \mu + \int_\Omega (\nabla w_1)^T a \nabla w_2\,, \quad (2.7)$$

then [66, Sec. 5] the optimal strategy of Player II is to select the distribution of ξ as that of a centered Gaussian field with covariance operator $\mathcal{L} = \frac{4}{\zeta^2}\mu\cdot - \operatorname{div}\big(a\nabla\cdot\big)$. This simply means that if $f \in H_0^1(\Omega)$ then $\int_\Omega f\xi$ is a centered Gaussian random variable of variance $\|f\|_\zeta^2$. Under that choice $u^{(k),\zeta}$ can be written as a linear combination of the measurements, i.e.

$$u^{(k),\zeta}(x) = \sum_{i\in\mathcal{I}^{(k)}} \psi_i^{(k),\zeta}(x)\int_\Omega u(y)\phi_i^{(k)}(y)\,dy \quad (2.8)$$

and the coefficients $\psi_i^{(k),\zeta}$ are deterministic functions and elementary gambles (namely, ζ-*gamblets* or *gamblets*) forming a basis for Player II's strategy ($\psi_i^{(k),\zeta}(x)$ is the best bet of Player II on the value of $u(x)$ given the information that $\int_\Omega u\phi_j^{(k)} = \delta_{i,j}$ for $j \in \mathcal{I}^{(k)}$). As shown in [62] (see also [61, 66]), gamblets are optimal recovery splines [54] characterized by optimal variational and recovery properties.

Theorem 2.4. *It holds true that (1) for $i \in \mathcal{I}^{(k)}$,*

$$\psi_i^{(k),\zeta} = \sum_{j\in\mathcal{I}^{(k)}} \Theta_{i,j}^{(k),-1}\mathcal{L}^{-1}\phi_j^{(k)} \quad (2.9)$$

where $\Theta^{(k),-1}$ is the inverse of the Gramian matrix $\Theta_{i,j}^{(k)} := \int_\Omega \phi_i^{(k)}\mathcal{L}^{-1}\phi_j^{(k)}$ (2) for $w \in \mathbb{R}^{\mathcal{I}^{(k)}}$, $\sum_{i\in\mathcal{I}^{(k)}} w_i\psi_i^{(k),\zeta}$ is the minimizer of $\|\psi\|_\zeta$ over all functions $\psi \in H_0^1(\Omega)$ such that $\int_\Omega \psi\phi_i^{(k)} = w_i$ for $i \in \mathcal{I}^{(k)}$ and (3) $u^{(k),\zeta}$ is the minimizer of $\|u-\psi\|_\zeta$ over all functions ψ in $\operatorname{span}\{\mathcal{L}^{-1}\phi_i^{(k)} \mid i \in \mathcal{I}^{(k)}\}$.

Furthermore $\psi_i^{(k),\zeta}$ decays exponentially fast away from the support of $\phi_i^{(k)}$ and this exponential decay can be used to localize the nested computation of gamblets. To simplify the presentation, we will write C any constant that depends only on $d, \Omega, \lambda_{\min}(a), \lambda_{\max}(a), \mu_{\min}, \mu_{\max}, \delta$ but not on ζ nor H (e.g., $2C\zeta H^2\lambda_{\max}(a)$ will be written $C\zeta H^2$).

Theorem 2.5. *Let $\phi_i^{(k)}$ be as in Construction 2.3. Let $\Omega_{i,n}^{(k)}$ be the union of subsets $\tau_j^{(k)}$ that are at distance at most nH from $\tau_i^{(k)}$. Let $\psi_i^{(k),\zeta,n}$ be the minimizer of $\|\psi\|_\zeta$ over all functions $\psi \in H_0^1(\Omega_{i,n}^{(k)})$ such that $\int_\Omega \psi\phi_j^{(k)} = \delta_{i,j}$ for $j \in \mathcal{I}^{(k)}$. We have $\|\psi_i^{(k),\zeta} - \psi_i^{(k),\zeta,n}\|_\zeta \leq C\|\psi_i^{(k),\zeta,0}\|_\zeta e^{-C^{-1}n}$*

Remark 2.6. *The optimal prior is Gaussian because [66, Sec. 5] of the linearity of the PDE and the quadratic nature of the loss function. For non linear PDEs or non quadratic loss functions, although optimal priors (which may not be Gaussian) could in principle be numerically approximated, such approximations could be severely impacted by stability issues as discussed in [67, 63, 68, 64].*

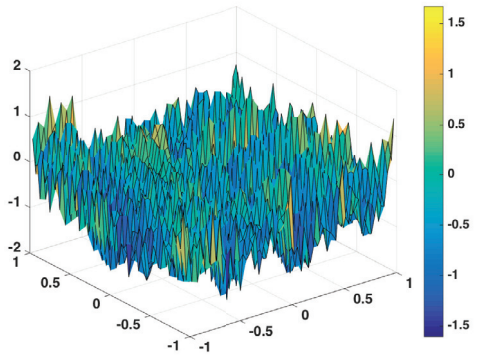


Figure 2: $a(x)$ in log scale.

Example 2.2. *For our numerical examples/illustrations, $d = 2$, $\Omega = (0,1)^2$ and $\mathcal{T}_h$ is a square grid of mesh size $h = (1+2^q)^{-1}$ with $r = 6$ and 64×64 interior nodes, a is piecewise constant on each square of $\mathcal{T}_h$ and given by*

$$a(x) = \Pi_{k=1}^r \big(1+0.5\cos(2^k\pi(\frac{i}{2^r+1}+\frac{j}{2^r+1}))\big)\big(1+0.5\sin(2^k\pi(\frac{j}{2^r+1}-3\frac{i}{2^r+1}))\big) \tag{2.10}$$

for $x \in [\frac{i}{2^r+1}, \frac{i+1}{2^r+1}) \times [\frac{j}{2^r+1}, \frac{j+1}{2^r+1})$ *as illustrated in Figure 2. We use continuous bilinear nodal basis elements* φ_i *spanned by* $\{1, x_1, x_2, x_1x_2\}$ *in each square of* $\mathcal{T}_h$*. Figure 3 then provides an illustration of gamblets for various values of* ζ*. Note that the generalized gamblet* $\psi_i^{(k),\zeta}$ *can be seen as a non-linear interpolation between a re-scaling of the measurement function* $\phi_i^{(k)}$ *(*$\zeta = 0$*) and the gamblets introduced in [62] (*$\zeta = \infty$*).*

2.5 Multiresolution decomposition

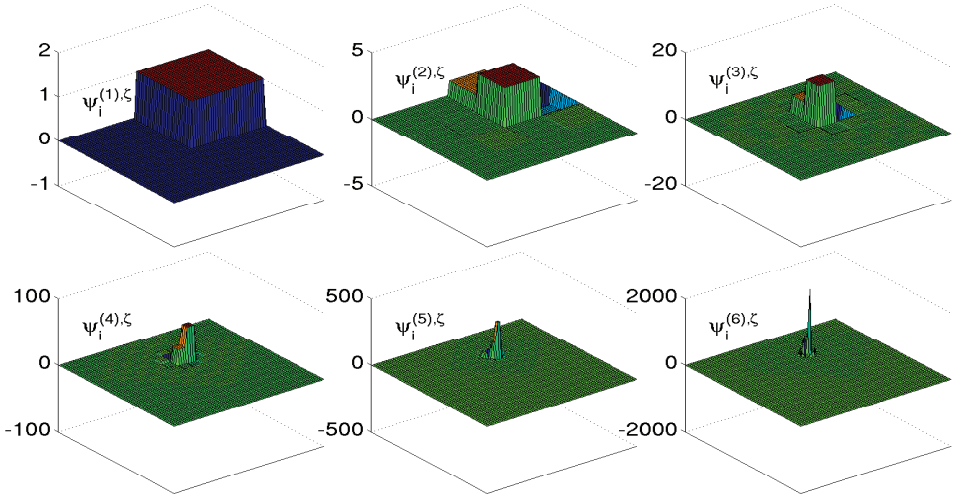
The nesting of the measurements function implies that of the gamblets, i.e. writing (for $k \in \{1, \dots, r\}$) $\mathfrak{V}^{(k),\zeta} := \operatorname{span}\{\psi_i^{(k),\zeta} \mid i \in \mathcal{I}^{(k)}\}$ we have (for $k \in \{1, \dots, r-1\}$) $\mathfrak{V}^{(k),\zeta} \subset \mathfrak{V}^{(k+1),\zeta}$ and $\psi_i^{(k),\zeta}(x) = \sum_{j \in \mathcal{I}_{k+1}} R_{i,j}^{(k),\zeta} \psi_j^{(k+1),\zeta}(x)$ where $R^{(k),\zeta}$ is the so called restriction/prolongation operator whose entry $R_{i,j}^{(k),\zeta}$ can be identified as, $R_{i,j}^{(k),\zeta} = \mathbb{E}\big[\int_\Omega v(y)\phi_j^{(k+1)}(y)\,dy \big| \int_\Omega v(y)\phi_l^{(k)}(y)\,dy = \delta_{i,l},\ l \in \mathcal{I}^{(k)}\big]$, i.e. the best bet of Player II on the value of $\int_\Omega u\phi_j^{(k+1),\zeta}$ given the information that $\int_\Omega u\phi_l^{(k),\zeta} = \delta_{i,l}$. With the identification of the restriction/prolongation operator one can use gamblets to couple scales in a multigrid algorithm but here we will instead use gamblets to induce a multiresolution decomposition of the solution space via orthogonalization process akin to the one used with wavelets.

Definition 2.7. *For* $k \in \{2, \dots, r\}$ *let* $\mathcal{J}^{(k)}$ *be a finite set of* k*-tuples of the form* $j = (j_1, \dots, j_k)$ *such that* $\{j^{(k-1)} \mid j \in \mathcal{J}^{(k)}\} = \mathcal{I}^{(k-1)}$ *and for* $i \in \mathcal{I}^{(k-1)}$*,* $\operatorname{Card}\{j \in \mathcal{J}^{(k)} \mid j^{(k-1)} = i\} = \operatorname{Card}\{s \in \mathcal{I}^{(k)} \mid s^{(k-1)} = i\} - 1$*.*

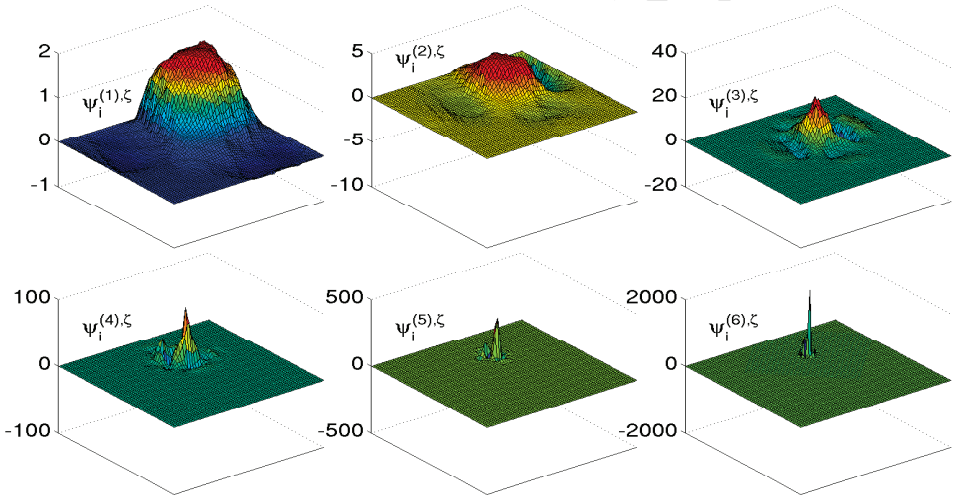
Definition 2.8. *Let* $W^{(k)}$ *be a* $\mathcal{J}^{(k)} \times \mathcal{I}^{(k)}$ *matrix such that: (1)* $\operatorname{Im}(W^{(k),T}) = \operatorname{Ker}(\pi^{(k-1,k)})$*, (2)* $W^{(k)}W^{(k),T} = J^{(k)}$ *where* $J^{(k)}$ *is the* $\mathcal{J}^{(k)} \times \mathcal{J}^{(k)}$ *identity matrix, (3)* $W_{j,i}^{(k)} = 0$ *for* $(j,i) \in \mathcal{J}^{(k)} \times \mathcal{I}^{(k)}$ *with* $j^{(k-1)} \neq i^{(k-1)}$*.*

When measurement functions are in Construction 2.3 then an example of $W^{(k)}$ is provided in Construction 2.10 based on the following lemma.

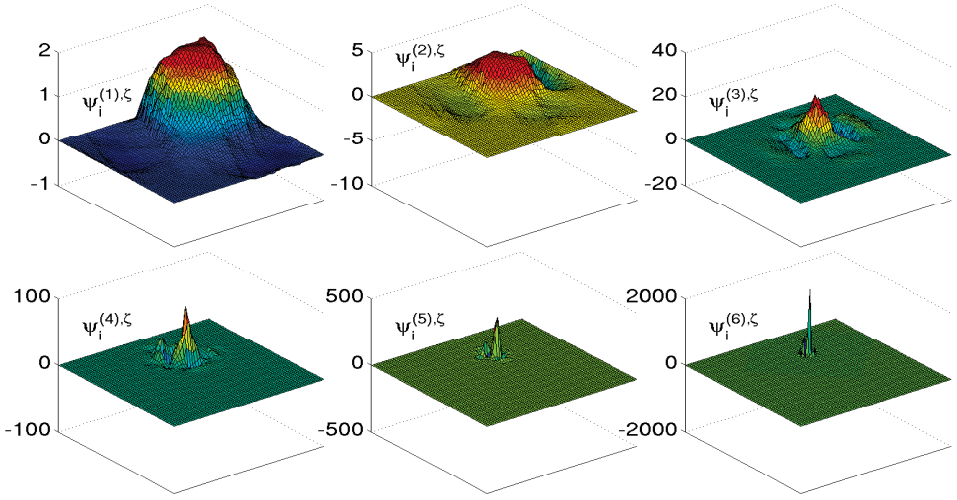
Lemma 2.9. *Let* $U^{(n)}$ *be the sequence of* $n \times n$ *matrices defined (1) for* $n = 2$ *by* $U_{1,\cdot}^{(2)} = (1, -1)$ *and* $U_{2,\cdot}^{(2)} = (1, 1)$ *and (2) iteratively for* $n \geq 2$ *by* $U_{i,j}^{(n+1)} = U_{i,j}^{(n)}$ *for* $1 \leq i, j \leq n$*,* $U_{n+1,j}^{(n+1)} = 1$ *for* $1 \leq j \leq n+1$*,* $U_{i,n+1}^{(n+1)} = 0$ *for* $1 \leq i \leq n-1$ *and* $U_{n,n+1}^{(n+1)} = -n$*. Then for* $n \geq 2$*, the rows of* $U^{(n)}$ *are orthogonal,* $U_{n,j}^{(n)} = 1$ *for* $1 \leq j \leq n$ *and we write* $\bar{U}^{(n)}$ *the corresponding orthonormal matrix obtained by renormalizing the rows of* $U^{(n)}$*.*



(a)



(b)



(c)

Figure 3: $\psi_i^{(k),\zeta}$ for (a) $\zeta = 10^{-6}$ (b) $\zeta = 1$ (c) $\zeta = 10^6$

Note that another possible choice for $U^{(n)}$ (than the one described in Lemma 2.9) is the discrete cosine transformation matrix.

Construction 2.10. *For $k \in \{2, \dots, r\}$, let $W^{(k)}$ be a $\mathcal{J}^{(k)} \times \mathcal{I}^{(k)}$ matrix such that: (1) $W^{(k)}_{j,i} = 0$ for $(j,i) \in \mathcal{J}^{(k)} \times \mathcal{I}^{(k)}$ with $j^{(k-1)} \neq i^{(k-1)}$, (2) for $s \in \mathcal{I}^{(k-1)}$ and $t \in \{1, \dots, n-1\}$ and $t' \in \{1, \dots, n\}$, $W^{(k)}_{(s,t),(s,t')} = \bar{U}^{(n)}_{t,t'}$ (where $\bar{U}^{(n)}$ is defined in Lemma 2.9 and $n = \mathrm{Card}\{i \in \mathcal{I}^{(k)} \mid i^{(k-1)} = s\}$).*

For $k \in \{2, \dots, r\}$ and $i \in \mathcal{J}^{(k)}$ let

$$\chi^{(k),\zeta}_i = \sum_{j \in \mathcal{I}^{(k)}} W^{(k)}_{i,j} \psi^{(k),\zeta}_j \tag{2.11}$$

and

$$\mathfrak{W}^{(k),\zeta} := \mathrm{span}\{\chi^{(k),\zeta}_i \mid i \in \mathcal{I}^{(k)}\} \tag{2.12}$$

For $k \in \{2, \dots, r\}$, write $\mathfrak{W}^{(k),\zeta} := \mathrm{span}\{\chi^{(k),\zeta}_i \mid i \in \mathcal{J}^{(k)}\}$. Write $\oplus_\zeta$ the orthogonal direct sum with respect to the scalar product $\langle \cdot, \cdot \rangle_\zeta$. The following theorem shows that $\mathfrak{W}^{(k),\zeta}$ is the orthogonal complement of $\mathfrak{V}^{(k),\zeta}$ in $\mathfrak{V}^{(k-1),\zeta}$ and this induces a multiresolution decomposition of the solution space.

Theorem 2.11. *It holds true that for $k \in \{2, \dots, r\}$, $\mathfrak{V}^{(k),\zeta} = \mathfrak{V}^{(k-1),\zeta} \oplus_\zeta \mathfrak{W}^{(k),\zeta}$ and, in particular*

$$\mathfrak{V}^{(r),\zeta} = \mathfrak{W}^{(1),\zeta} \oplus_\zeta \mathfrak{W}^{(2),\zeta} \oplus_\zeta \cdots \oplus_\zeta \mathfrak{W}^{(r),\zeta}, \tag{2.13}$$

where $\mathfrak{W}^{(1),\zeta} = \mathfrak{V}^{(1),\zeta}$. Furthermore, $u^{(1)}$ is the finite-element solution of (2.5) in $\mathfrak{V}^{(1),\zeta}$ and for $k \in \{2, \dots, r\}$, $u^{(k),\zeta} - u^{(k-1),\zeta}$ is the finite element solution of (2.5) in $\mathfrak{W}^{(k),\zeta}$.

Note that since the spaces $\mathfrak{W}^{(k),\zeta}$ for $k \in \{1, \dots, r\}$ are orthogonal with each other, the corresponding finite-element subband solutions $u^{(1),\zeta}$ and $u^{(k),\zeta} - u^{(k-1),\zeta}$ for $k \in \{2, \dots, r\}$ can be computed independently. Figure 5 provides an illustration of the subband solutions for the solution $u(x)$ of equation (2.1) with $g(x) = \sin(\pi x_1)\cos(\pi x_2)$.

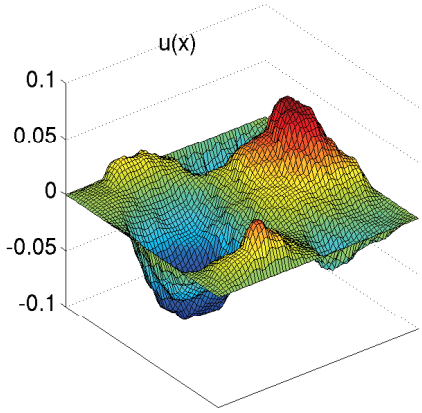


Figure 4: Solution $u(x)$ of equation (2.1) with $g(x) = \sin(\pi x_1)\cos(\pi x_2)$.

Remark 2.12. *An analogy could be made between gamblets and Lax Pairs [49] where the solution space is also decomposed in a way that involves the dynamics of the PDE itself. Here the pairs* $\phi_i^{(k)}$ *and* $\phi_i^{(k),\zeta}$ *form a biorthogonal system [27] in the sense that* $\int_\Omega \phi_i^{(k)} \psi_j^{(k),\zeta} = \delta_{i,j}$ *for* $i,j \in \mathcal{I}^{(k)}$ *and the* $\langle \cdot,\cdot \rangle_\zeta$*-orthogonal projection of* $u \in H_0^1(\Omega)$ *onto* $\mathfrak{V}^{(k),\zeta}$ *is* $\sum_{i\in\mathcal{I}^{(k)}} \psi_i^{(k)} \int_\Omega \phi_i^{(k)} u$*. As discussed in [66] gamblets are also optimal recovery splines in the sense of Micchelli and Rivlin [54] and as a consequence have optimal recovery properties [66].*

2.6 Uniformly bounded condition numbers

Let $A^{(k),\zeta}$ and $B^{(k),\zeta}$ be the stiffness matrices of finite element approximation of the (2.5) in $\mathfrak{V}^{(k),\zeta}$ and $\mathfrak{W}^{(k),\zeta}$, i.e. $A_{i,j}^{(k),\zeta} := \langle \psi_i^{(k)}, \psi_j^{(k)} \rangle_\zeta$ for $k \in \{1,\dots,r\}$ and $i,j \in \mathcal{I}^{(k)}$ and $B_{i,j}^{(k),\zeta} := \langle \chi_i^{(k),\zeta}, \chi_j^{(k),\zeta} \rangle_\zeta$ for $k \in \{2,\dots,r\}$ and $i,j \in \mathcal{J}^{(k)}$. The following theorem shows that the multiresolution decomposition of Theorem (2.11) looks like an eigenspace decomposition in the sense that those subspaces are orthogonal with respect to the scalar product $\langle \cdot,\cdot \rangle$ and the condition numbers of the matrices $B^{(k),\zeta}$ are uniformly bounded (the subspaces are not orthogonal in $L^2(\Omega)$ so (2.13) is not an exact eigenspace decomposition). For a given matrix M, write $\mathrm{Cond}(M) := \sqrt{\lambda_{\max}(M^T M)}/\sqrt{\lambda_{\min}(M^T M)}$ its condition number. If M

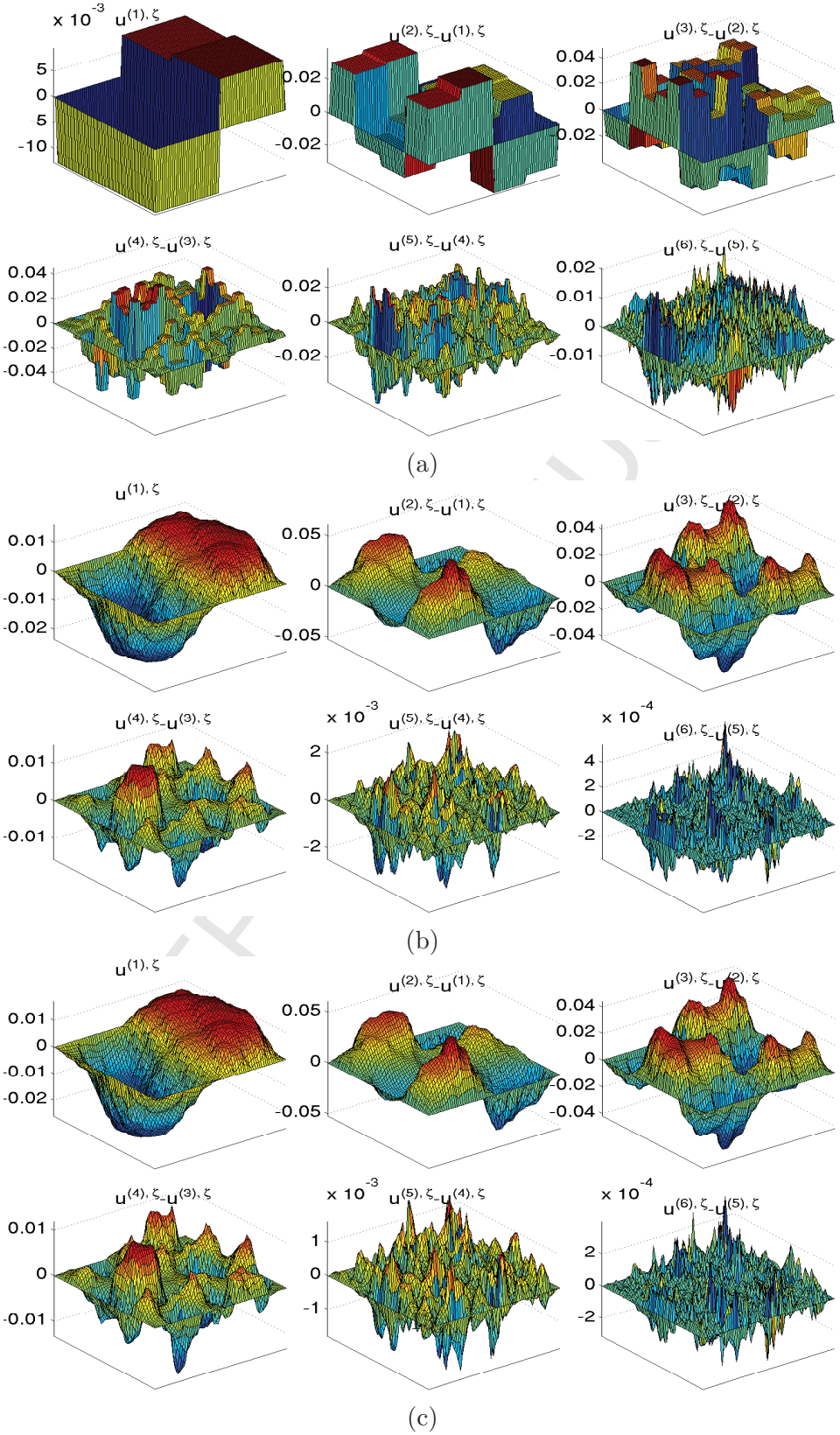


Figure 5: Multiresolution decomposition of u for $\zeta = 10^{-6}$, $\zeta = 1$ and $\zeta = 10^6$

is symmetric write $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ its minimal and maximal eigenvalues.

Theorem 2.13. *Let the $\phi_i^{(k)}$ be as in Construction 2.3. For $\zeta \in (0, \infty]$, $\mathrm{Cond}(A^{(1),\zeta}) \leq CH^{-2}$, and $\mathrm{Cond}(B^{(k),\zeta}) \leq CH^{-2}$ for $k \in \{2, \ldots, r\}$. Furthermore, for $\zeta = \infty$ and $k \in \{1, \ldots, r\}$, $\frac{1}{C} \leq \lambda_{\min}(A^{(k),\infty})$ and $\lambda_{\max}(A^{(k),\infty}) \leq CH^{-2k}$. For $\zeta = \infty$ and $k \in \{2, \ldots, r\}$, $\frac{1}{C}H^{-2(k-1)} \leq \lambda_{\min}(B^{(k),\infty})$ and $\lambda_{\max}(B^{(k),\infty}) \leq CH^{-2k}$.*

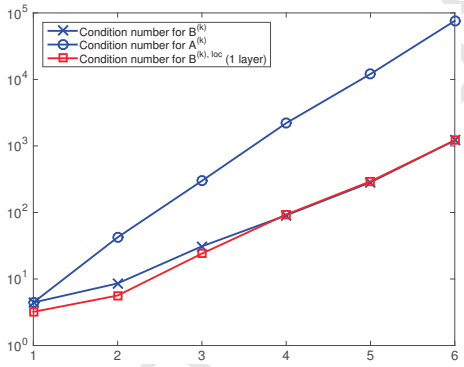


Figure 6: Condition numbers of $B^{(k)}$ ($\zeta = \infty$) for $k = 1, \ldots, 6$ and $a(x)$ defined as in (2.10).

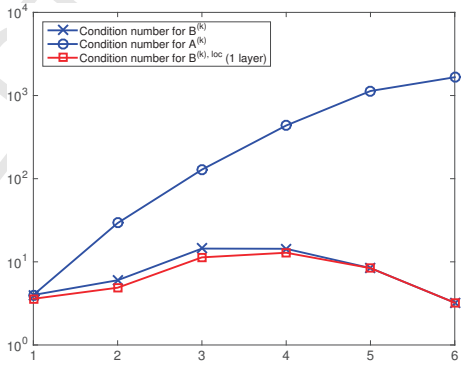


Figure 7: Condition numbers of $B^{(k)}$ ($\zeta = \infty$) for $k = 1, \ldots, 6$ and $a(x) = I_d$ (the Laplacian).

Example 2.3. *To simplify notations we will write $A^{(1),\zeta}$ as $B^{(1),\zeta}$ and omit the superscript ζ when $\zeta = \infty$. Figures 6 and 7 provide the condition num-*

bers of $A^{(k)}$ *and* $B^{(k)}$ *for* $k = 1, \ldots, 6$ *for* $a(x)$ *defined as in* (2.10) *and* $a(x) = I_d$ *(the Laplacian). Note that these condition numbers do depend on the contrast of* a*. Figures 8 and 9 illustrate the ranges of the eigenvalues of the PDE in* $\mathfrak{V}$ *and in each subband* $\mathfrak{W}^{(k)}$*, for* $a(x)$ *defined as in* (2.10) *and* $a(x) = I_d$ *(the Laplacian), i.e. the figures are illustrations of the intervals* $\left[\inf_{\psi\in\mathfrak{V}} \frac{\|\psi\|_a}{\|\psi\|_{L^2(\Omega)}}, \sup_{\psi\in\mathfrak{V}} \frac{\|\psi\|_a}{\|\psi\|_{L^2(\Omega)}}\right]$ *and* $\left[\inf_{\psi\in\mathfrak{W}^{(k)}} \frac{\|\psi\|_a}{\|\psi\|_{L^2(\Omega)}}, \sup_{\psi\in\mathfrak{W}^{(k)}} \frac{\|\psi\|_a}{\|\psi\|_{L^2(\Omega)}}\right]$ *where we write* $\|\psi\|_a := \|\psi\|_\zeta$ *for* $\zeta = \infty$*. Note that the eigenvalues of* $B^{(k)}$ *cover only subintervals of spectrum of the discretized operator, which corresponds to a multi-resolution decomposition.*

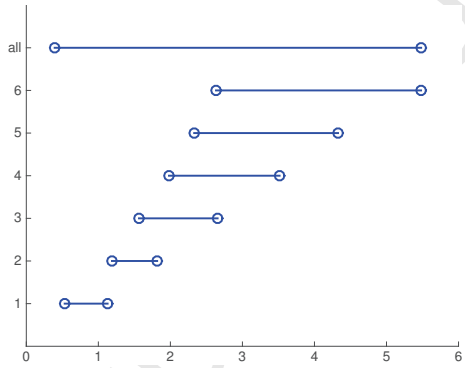


Figure 8: Ranges of eigenvalues in $\mathfrak{V}$ and $\mathfrak{W}^{(k)}$ ($\zeta = \infty$) for $k = 1, \ldots, 6$ and $a(x)$ defined as in (2.10).

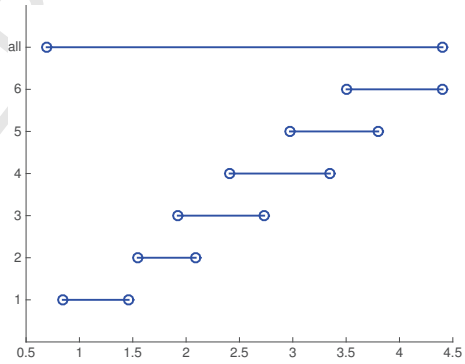


Figure 9: Ranges of eigenvalues in $\mathfrak{V}$ and $\mathfrak{W}^{(k)}$ ($\zeta = \infty$) for $k = 1, \ldots, 6$ and $a(x) = I_d$ (the Laplacian)

2.7 Algorithms

2.7.1 The gamblet transform

We will now describe the gamblet transform for the discrete operator obtained from the numerical approximation of (2.1). Consider the finite-element solution of (2.1) over a basis $(\varphi_i)_{i\in\mathcal{N}}$ of fine-scale elements. To facilitate the presentation, assume that $(\varphi_i)_{i\in\mathcal{N}}$ is obtained from $\mathcal{T}_h$, a regular fine mesh discretization of Ω of resolution h with $0 < h \ll 1$. Let $\mathcal{N}$ be the set of interior nodes z_i and $N = |\mathcal{N}|$ be the number of interior nodes ($N = \mathcal{O}(h^{-d})$) of $\mathcal{T}_h$. Write $(\varphi_i)_{i\in\mathcal{N}}$ a set of conforming nodal basis elements (of $H^1_0(\Omega)$) constructed from $\mathcal{T}_h$ such that for each $i \in \mathcal{N}$, $\text{support}(\varphi_i) \subset B(z_i, C_0 h)$, for $y \in \mathbb{R}^N$,

$$\underline{\gamma} h^d |y|^2 \leq \| \sum_{i\in\mathcal{N}} y_i \varphi_i \|^2_{L^2(\Omega)} \leq \bar{\gamma} h^d |y|^2 \tag{2.14}$$

and

$$\|\nabla v\|_{L^2(\Omega)} \leq C_1 h^{-1} \|v\|_{L^2(\Omega)} \tag{2.15}$$

for $v \in \text{span}\{\varphi_i \mid i \in \mathcal{N}\}$, for some constants $\underline{\gamma}, \bar{\gamma}, C_0, C_1 \approx \mathcal{O}(1)$.

In addition to properties (2.14) and (2.15) we assume that $\mathcal{T}_h$ is such that: (1) $h = H^r$ and (2) each set $\tau_i^{(r)}$ ($i \in \mathcal{I}^{(r)}$) contains one and only one interior node of $\mathcal{T}_h$. Using this one to one correspondence we use the elements of $\mathcal{I}^r$ to relabel the interior nodes z_i of $\mathcal{T}_h$ and their respective nodal elements φ_i.

Write $\mathfrak{V} := \text{span}\{\varphi_i \mid i \in \mathcal{I}\}$. Given the matrices $\pi^{(k,k+1)}$ defined as in constructions 2.2 and 2.4, and the matrices $W^{(k)}$ obtained as in Definition 2.8, Algorithm 1 computes elements of $\mathfrak{V}$ corresponding discrete gamblets $(\psi_i^{(k),\zeta})_{i\in\mathcal{I}^{(k)}}$ and $\{\chi_i^{(k),\zeta}\}_{i\in\mathcal{J}^{(k)}}$. As in Theorem 2.11, these discrete gamblets induce the following $\langle\cdot,\cdot\rangle_\zeta$ orthogonal decomposition of the solution space, corresponding to a diagonalization of the stiffness matrix $A_{i,j} = \langle\varphi_i, \varphi_j\rangle_z$ into blocks of uniformly bounded condition numbers (which can be approximated by truncated blocks thanks to the exponential decay of gamblets).

$$\mathfrak{V} = \mathfrak{V}^{(1),\zeta} \oplus_\zeta \mathfrak{W}^{(2),\zeta} \oplus_\zeta \cdots \oplus_\zeta \mathfrak{W}^{(r),\zeta}, \tag{2.16}$$

Observe that the measurement functions $\phi_i^{(k)}$ do not appear explicitly in Algorithm 1 (which depends only on the interpolation matrices $\pi^{(k,k+1)}$ defined in Constructions 2.2 and 2.4). Note that at the finest scale, level r gamblets $\psi_i^{(r),\zeta}$ are simply the basis elements φ_i used to discretize the PDE

(2.1) (as discussed in [66], writing $\bar{M}$ the mass matrix $\bar{M}^{\varphi}_{i,j} = \int_\Omega \varphi_i \varphi_j$, selecting $\psi_i^{(r),\zeta} = \varphi_i$ is equivalent to using $\phi_i^{(r)} = \sum_{j\in\mathcal{I}^{(r)}} \bar{M}^{-1}_{i,j}\varphi_j$ as level r measurement functions and defining $\phi_i^{(k)}$ via aggregation as in (2.4)).

Algorithm 1 Gamblet Transform.

1: For $i,j \in \mathcal{I}^{(r)}$, $A^{\varphi,\zeta}_{i,j} = \langle \varphi_i, \varphi_j \rangle_\zeta$ // Stiffness matrix
2: For $i \in \mathcal{I}^{(r)}$, $\psi_i^{(r),\zeta} = \varphi_i$ // Level r gamblets
3: For $i,j \in \mathcal{I}^{(r)}$, $A^{(r),\zeta}_{i,j} = \langle \psi_i^{(r),\zeta}, \psi_j^{(r),\zeta} \rangle_\zeta$ // $A^{(r),\zeta} = A^{\varphi,\zeta}$
4: **for** $k = r$ to 2 **do**
5: For $i \in \mathcal{J}^{(k)}$, $\chi_i^{(k),\zeta} = \sum_{j\in\mathcal{I}^{(k)}} W^{(k)}_{i,j}\psi_j^{(k),\zeta}$ // Level k, χ gamblets
6: $B^{(k),\zeta} = W^{(k)}A^{(k),\zeta}W^{(k),T}$ // $B^{(k),\zeta}_{i,j} = \langle \chi_i^{(k),\zeta}, \chi_j^{(k),\zeta} \rangle_\zeta$
7: $D^{(k,k-1),\zeta} = -B^{(k),\zeta,-1}W^{(k)}A^{(k),\zeta}\bar{\pi}^{(k,k-1)}$ // $B^{(k),\zeta,-1}$ =matrix inverse of $B^{(k),\zeta}$
8: $R^{(k-1,k),\zeta} = \bar{\pi}^{(k-1,k)} + D^{(k-1,k),\zeta}W^{(k)}$ // Interpolation/restriction operator
9: For $i \in \mathcal{I}^{(k-1)}$, $\psi_i^{(k-1),\zeta} = \sum_{j\in\mathcal{I}^{(k)}} R^{(k-1,k),\zeta}_{i,j}\psi_j^{(k),\zeta}$ // Level $k-1$, ψ gamblets
10: $A^{(k-1),\zeta} = R^{(k-1,k),\zeta}A^{(k),\zeta}R^{(k,k-1),\zeta}$ // $A^{(k-1),\zeta}_{i,j} = \langle \psi_i^{(k-1),\zeta}, \psi_j^{(k-1),\zeta} \rangle_\zeta$
11: **end for**

2.7.2 Linear solve with gamblets

Given $g = \sum_{i\in\mathcal{N}} g_i\varphi_i$, Algorithm 2 computes $u \in \mathrm{span}\{\varphi_i \mid i \in \mathcal{N}\}$ such that,

$$\langle \varphi_j, u \rangle_\zeta = \int_\Omega \varphi_j g, \text{ for all } j \in \mathcal{N} \tag{2.17}$$

Algorithm 2 is exact and $u = u^{(1),\zeta} + (u^{(2),\zeta} - u^{(1),\zeta}) + \cdots + (u^{(r),\zeta} - u^{(r-1),\zeta})$ obtained in Line 9 is the orthogonal decomposition of the solution u of (2.17) over $\mathfrak{V} = \mathfrak{V}^{(1),\zeta} \oplus_\zeta \mathfrak{W}^{(2),\zeta} \oplus_\zeta \cdots \oplus_\zeta \mathfrak{W}^{(r),\zeta}$.

2.8 Fast gamblet transform

Algorithms 1 and 2 can be modified to operate in linear complexity. This near linear complexity is possible thanks to three main properties, (1) Nesting: level k gamblets and stiffness matrices can be computed from level $k+1$

Algorithm 2 Linear solve with exact gamblets.

For $i \in \mathcal{I}^{(r)}$, $g_i^{(r),\zeta} = g_i$ // $g_i^{(r),\zeta} = \int_\Omega \psi_i^{(r),\zeta} g$ with $g = \sum_{i\in\mathcal{I}^{(r)}} g_i \varphi_i$
for $k = r$ to 2 **do**
  $w^{(k),\zeta} = B^{(k),\zeta,-1} W^{(k)} g^{(k),\zeta}$
  $u^{(k),\zeta} - u^{(k-1),\zeta} = \sum_{i\in\mathcal{J}^{(k)}} w_i^{(k),\zeta} \chi_i^{(k),\zeta}$
  $g^{(k-1),\zeta} = R^{(k-1,k),\zeta} g^{(k),\zeta}$
end for
$U^{(1),\zeta} = A^{(1),\zeta,-1} g^{(1),\zeta}$
$u^{(1),\zeta} = \sum_{i\in\mathcal{I}^{(1)}} U_i^{(1),\zeta} \psi_i^{(1),\zeta}$
$u = u^{(1),\zeta} + (u^{(2),\zeta} - u^{(1),\zeta}) + \cdots + (u^{(r),\zeta} - u^{(r-1),\zeta})$

gamblets and stiffness matrices; (2) Bounded condition numbers: It follows from Theorem 2.13 that the linear systems involved in Algorithms 1 and 2 have uniformly bounded condition numbers; (3) Localization: gamblets can be localized as a function of the desired accuracy. The resulting modified algorithms are 3 and 4.

Algorithm 3 Localized Gamblets.

For $i \in \mathcal{I}^{(r)}$, $\psi_i^{(r),\zeta,\mathrm{loc}} = \varphi_i$
For $i, j \in \mathcal{I}^{(r)}$, $A_{i,j}^{(r),\zeta,\mathrm{loc}} = \langle \psi_i^{(r),\zeta,\mathrm{loc}}, \psi_j^{(r),\zeta,\mathrm{loc}} \rangle_\zeta$
for $k = r$ to 2 **do**
  $B^{(k),\zeta,\mathrm{loc}} = W^{(k)} A^{(k),\zeta,\mathrm{loc}} W^{(k),T}$
  For $i \in \mathcal{J}^{(k)}$, $\chi_i^{(k),\zeta,\mathrm{loc}} = \sum_{j\in\mathcal{I}^{(k)}} W_{i,j}^{(k)} \psi_j^{(k),\zeta,\mathrm{loc}}$
  $\mathrm{Inv}(B^{(k),\zeta,\mathrm{loc}} D^{(k,k-1),\zeta,\mathrm{loc}} = -W^{(k)} A^{(k),\zeta,\mathrm{loc}} \bar{\pi}^{(k,k-1)}, \rho_{k-1})$ // Def. 2.14, Thm. 2.18
  $R^{(k-1,k),\zeta,\mathrm{loc}} = \bar{\pi}^{(k-1,k)} + D^{(k-1,k),\zeta,\mathrm{loc}} W^{(k)}$ // Def. 2.14
  $A^{(k-1),\zeta,\mathrm{loc}} = R^{(k-1,k),\zeta,\mathrm{loc}} A^{(k),\zeta,\mathrm{loc}} R^{(k,k-1),\zeta,\mathrm{loc}}$
  For $i \in \mathcal{I}^{(k-1)}$, $\psi_i^{(k-1),\zeta,\mathrm{loc}} = \sum_{j\in\mathcal{I}^{(k)}} R_{i,j}^{(k-1,k),\zeta,\mathrm{loc}} \psi_j^{(k),\zeta,\mathrm{loc}}$
end for

Algorithm 3 achieves $\mathcal{O}\big(N \ln^{3d} N\big)$ complexity in computing approximate gamblets (sufficient to achieve a given level of accuracy). This fast algorithm is obtained by localizing/truncating the linear systems corresponding to Line refline11 in Algorithm 1. The approximation error induced by these localization/truncation steps is controlled by the exponential decay of gamblets $\psi_i^{(k),\zeta}$ and $\chi_i^{(k),\zeta}$ and the uniform bound on the condition numbers of the

matrices $B^{(k),\zeta}$ and $A^{(1),\zeta}$. We define these localization/truncation steps as follows. For $k \in \{1,\dots,q\}$ and $i \in \mathcal{I}^{(k)}$ define i^ρ as the subset of indices $j \in \mathcal{I}^{(k)}$ whose corresponding subdomains $\tau_j^{(k)}$ are at distance at most $H_k\rho$ from $\tau_i^{(k)}$.

Note that level r gamblets $\psi_i^{(r),\zeta,\mathrm{loc}}$ are simply the finite-elements φ_i used to discretize the operator. (Line 1 of Algorithm 3). Line 6 of Algorithm 3 is defined as follows.

Definition 2.14. *Let $k \in \{2,\dots,r\}$ and B be the positive definite $\mathcal{J}^{(k)} \times \mathcal{J}^{(k)}$ matrix $B^{(k),\zeta,\mathrm{loc}}$ computed in Line 4 of Algorithm 3. For $i \in \mathcal{I}^{(k-1)}$, let $\rho = \rho_{k-1}$ and let i^χ be the subset of indices $j \in \mathcal{J}^{(k)}$ such that $j^{(k-1)} \in i^\rho$ (recall that if $j = (j_1,\dots,j_k) \in \mathcal{J}^{(k)}$ then $j^{(k-1)} := (j_1,\dots,j_{k-1}) \in \mathcal{I}^{(k-1)}$). $B^{(i,\rho)}$ be the $i^\chi \times i^\chi$ matrix defined by $B^{(i,\rho)}_{l,j} = B_{l,j}$ for $l,j \in i^\chi$. Let $b^{(i,\rho)}$ be the $|i^\chi|$-dimensional vector defined by $b^{(i,\rho)}_j = -(W^{(k)}A^{(k),\zeta,\mathrm{loc}}\bar{\pi}^{(k,k-1)})_{j,i}$ for $j \in i^\chi$. Let $y^{(i,\rho)}$ be the $|i^\chi|$-dimensional vector solution of $B^{(i,\rho)}y^{(i,\rho)} = b^{(i,\rho)}$. We define the solution $D^{(k,k-1),\zeta,\mathrm{loc}}$ of the localized linear system $\mathrm{Inv}(B^{(k),\zeta,\mathrm{loc}}D^{(k,k-1),\zeta,\mathrm{loc}} = -W^{(k)}A^{(k),\zeta,\mathrm{loc}}\bar{\pi}^{(k,k-1)}, \rho_{k-1})$ as the $\mathcal{J}^{(k)}\times\mathcal{I}^{(k-1)}$ sparse matrix given by $D^{(k,k-1),\zeta,\mathrm{loc}}_{j,i} = 0$ for $j \notin i^\chi$ and $D^{(k,k-1),\zeta,\mathrm{loc}}_{j,i} = y^{(i,\rho)}_j$ for $j \in i^\chi$. $D^{(k-1,k),\zeta,\mathrm{loc}}$ (Line 7 of Algorithm 3) is then defined as the transpose of $D^{(k,k-1),\zeta,\mathrm{loc}}$.*

Remark 2.15. *Definition 2.14 (Line 4 of Algorithm 3) is equivalent to localizing the computation of each gamblet $\psi_i^{(k-1),\zeta}$ to a subdomain of size $H_{k-1}\rho_{k-1}$, i.e., the gamblet $\psi_i^{(k-1),\zeta,\mathrm{loc}}$ computed in Line 9 of Algorithm 3 is the solution of (1) the problem of finding ψ in the affine space $\sum_{j\in\mathcal{I}^{(k)}}\bar{\pi}^{(k-1,k)}_{i,j}\psi^{(k),\zeta,\mathrm{loc}}_j +$ $\mathrm{span}\{\chi^{(k),\zeta,\mathrm{loc}}_j \mid j^{(k-1)} \in i^{\rho_{k-1}}\}$ such that ψ is $\langle\cdot,\cdot\rangle_\zeta$ orthogonal to $\mathrm{span}\{\chi^{(k),\zeta,\mathrm{loc}}_j \mid j^{(k-1)} \in i^{\rho_{k-1}}\}$, and (2) the problem of minimizing $\|\psi\|_\zeta$ in $\mathrm{span}\{\psi^{(k),\zeta,\mathrm{loc}}_l \mid l^{(k-1)} \in i^{\rho_{k-1}}\}$ subject to constraints $\int_\Omega \phi^{(k-1)}_j\psi = \delta_{i,j}$ for $j \in i^{\rho_{k-1}}$.*

We will (occasionally) write H_k for H^k to emphasize that, as in [62], the essentially property of the sequence H_k is that H_k/H_{k+1} remains uniformly bounded away from 1 and ∞.

To simplify the presentation, we will write C any constant that depends only on $d,\Omega,\lambda_{\min}(a),\lambda_{\max}(a),\delta,\underline{\gamma},\bar{\gamma},C_0,C_1,\mu_{\min},\mu_{\max},\delta$ but not on h,ζ nor H (e.g., $2C\zeta H^2\lambda_{\max}(a)$ will be written $C\zeta H^2$).

The following theorem shows that the condition numbers of the localized stiffness matrices $B^{(k),\zeta,\mathrm{loc}}$ remain uniformly bounded provided that the computation of level k gamblets is localized to subdomains of size $H_k \ln\frac{1}{H_k}$.

Theorem 2.16. *Let $\varepsilon \in (0,1)$. Assume that*

1. *$\rho_k \geq C\big((1+\frac{1}{\ln(1/H)})\ln\frac{1}{H_k}+\ln\frac{1}{\varepsilon}\big)$ for $k\in\{1,\dots,r\}$.*
2. *For $k\in\{2,\dots,r\}$ and each $i\in\mathcal{I}^{(k-1)}$, the localized linear system $B^{(i,\rho)}y=b$ of Definition 2.14 and Line 6 of Algorithm 3 is solved up to accuracy $|y-y^{\mathrm{ap}}|_{B^{(i,\rho)}}\leq C^{-1}H^{3-k+kd/2}\varepsilon/k^2$ (using the notation $|e|_A^2:=e^TAe$, and writing y^{ap} the approximation of y).*

Then it holds true that $\mathrm{Cond}(A^{(1),\zeta,\mathrm{loc}})\leq CH^{-2}$ *and for* $k\in\{2,\dots,r\}$, $\mathrm{Cond}(B^{(k),\zeta,\mathrm{loc}})\leq CH^{-2}$. *Furthermore for* $k\in\{1,\dots,r\}$ *and* $\zeta=\infty$, $\frac{1}{C}\leq\lambda_{\min}(A^{(k),\zeta,\mathrm{loc}})$ *and* $\lambda_{\max}(A^{(k),\zeta,\mathrm{loc}})\leq CH^{-2k}$, *and for* $k\in\{2,\dots,r\}$ *and* $\zeta=\infty$, $\frac{1}{C}H^{-2(k-1)}\leq\lambda_{\min}(B^{(k),\zeta,\mathrm{loc}})$ *and* $\lambda_{\max}(B^{(k),\zeta,\mathrm{loc}})\leq CH^{-2k}$. *Additionally the functions* $(\psi_i^{(1),\zeta,\mathrm{loc}})_{i\in\mathcal{I}^{(1)}}$ *and* $(\chi_i^{(k),\zeta,\mathrm{loc}})_{k\in\{2,\dots,r\},i\in\mathcal{J}^{(k)}}$ *are linearly independent and form a basis of* $\mathfrak{V}$.

We now present Algorithm 4, which computes an approximation of the solution of (2.17) using localized gamblets (up to ε accuracy in $H_0^1(\Omega)$-norm).

Algorithm 4 Linear solve with localized gamblets.

For $i\in\mathcal{I}^{(r)}$, $g_i^{(r),\zeta,\mathrm{loc}}=g_i$ // $g=\sum_{i\in\mathcal{I}^{(r)}}g_i\varphi_i$
for $k=r$ to 2 **do**
  $w^{(k),\zeta,\mathrm{loc}}=B^{(k),\zeta,\mathrm{loc},-1}W^{(k)}g^{(k),\zeta,\mathrm{loc}}$
  $u^{(k),\zeta,\mathrm{loc}}-u^{(k-1),\zeta,\mathrm{loc}}=\sum_{i\in\mathcal{J}^{(k)}}w_i^{(k),\zeta,\mathrm{loc}}\chi_i^{(k),\zeta,\mathrm{loc}}$
  $g^{(k-1),\zeta,\mathrm{loc}}=R^{(k-1,k),\zeta,\mathrm{loc}}g^{(k),\zeta,\mathrm{loc}}$
end for
$U^{(1),\zeta,\mathrm{loc}}=A^{(1),\zeta,\mathrm{loc},-1}g^{(1),\zeta,\mathrm{loc}}$
$u^{(1),\zeta,\mathrm{loc}}=\sum_{i\in\mathcal{I}^{(1)}}U_i^{(1),\zeta,\mathrm{loc}}\psi_i^{(1),\zeta,\mathrm{loc}}$
$u^{\mathrm{loc}}=u^{(1),\zeta,\mathrm{loc}}+(u^{(2),\zeta,\mathrm{loc}}-u^{(1),\zeta,\mathrm{loc}})+\cdots+(u^{(r),\zeta,\mathrm{loc}}-u^{(r-1),\zeta,\mathrm{loc}})$

Theorem 2.17. *Let u be the solution of the discrete system (2.17). Let $u^{(1),\zeta,\mathrm{loc}}$, $u^{(k),\zeta,\mathrm{loc}}-u^{(k-1),\zeta,\mathrm{loc}}$, u^{loc}, $A^{(k),\zeta,\mathrm{loc}}$ and $B^{(k),\zeta,\mathrm{loc}}$ be the outputs of algorithms 3 and 4. Let $u^{(1),\zeta}$ and $u^{(k),\zeta}-u^{(k-1),\zeta}$ be the outputs of Algorithm 2. For $k\in\{2,\dots,r\}$, write $u^{(k),\zeta,\mathrm{loc}}:=u^{(1),\zeta,\mathrm{loc}}+\sum_{j=2}^k(u^{(j),\zeta,\mathrm{loc}}-u^{(j-1),\zeta,\mathrm{loc}})$. Let $\varepsilon\in(0,1)$, it holds true that if $\rho_k\geq C\big((1+\frac{1}{\ln(1/H)})\ln\frac{1}{H_k}+\ln\frac{1}{\varepsilon}\big)$ for $k\in\{1,\dots,r\}$ then*

1. *for* $k \in \{1, \ldots, r-1\}$ *we have* $\|u^{(k),\zeta} - u^{(k),\zeta,\mathrm{loc}}\|_\zeta \leq \varepsilon \|g\|_{H^{-1}(\Omega)}$ *and* $\|u^{(k),\zeta} - u^{(k),\zeta,\mathrm{loc}}\|_\zeta \leq C(H_k + \varepsilon)\|g\|_{L^2(\Omega)}$
2. $\|u^{(k),\zeta} - u^{(k-1),\zeta} - (u^{(k),\zeta,\mathrm{loc}} - u^{(k-1),\zeta,\mathrm{loc}})\|_\zeta \leq \frac{\varepsilon}{2k^2}\|g\|_{H^{-1}(\Omega)}$.
3. *Furthermore,* $\|u - u^{\mathrm{loc}}\|_\zeta \leq \varepsilon \|g\|_{H^{-1}(\Omega)}$.

Theorem 2.18. *The results of Theorem 2.17 remain true if*

1. $\rho_k \geq C\big((1 + \frac{1}{\ln(1/H)}) \ln \frac{1}{H_k} + \ln \frac{1}{\varepsilon}\big)$ *for* $k \in \{1, \ldots, r\}$
2. *For* $k \in \{2, \ldots, r\}$ *and each* $i \in \mathcal{I}^{(k-1)}$, *the localized linear system* $B^{(i,\rho)}y = b$ *of Definition 2.14 and Line 6 of Algorithm 3 is solved up to accuracy* $|y - y^{\mathrm{ap}}|_{B^{(i,\rho)}} \leq C^{-1}H^{3-k+kd/2}\varepsilon/k^2$ *(using the notation* $|e|_A^2 := e^T A e$, *and writing* y^{ap} *the approximation of* y*).*
3. *For* $k \in \{2, \ldots, r\}$ *the linear system* $B^{(k),\zeta,\mathrm{loc}}y = W^{(k)}g^{(k),\zeta,\mathrm{loc}}$ *of Line 3 of Algorithm 4 is solved up to accuracy* $|y - y^{\mathrm{ap}}|_{B^{(k),\zeta,\mathrm{loc}}} \leq \varepsilon\|g\|_{H^{-1}(\Omega)}/(2r)$.

Observe that theorems 2.17 and 2.18 imply that (1) the complexity of Algorithm 3 is $\mathcal{O}\big(N\big(\ln \max(\frac{1}{\varepsilon}, N^{\frac{1}{d}})\big)^{3d}\big)$ (2) the complexity of Algorithm 4 is $\mathcal{O}\big(N\big(\ln \max(\frac{1}{\varepsilon}, N^{\frac{1}{d}})\big)^{d} \ln \frac{1}{\varepsilon}\big)$. Therefore if ε corresponds to a grid size accuracy $N^{-1/d}$ (in H^1-norm) then the complexity of Algorithm 3 is $\mathcal{O}(N \ln^{3d} N)$ and that of Algorithm 4 is $\mathcal{O}(N \ln^{d+1} N)$. In Figure 10, we show the elapsed time of localized Gamblet transform and localized gamblet linear solve for fixed $\rho_k = 3$ with respect to the degrees of freedom N. Although our implementation is in Matlab and it is not optimal in terms of efficiency, we can still observe close to linear complexity.

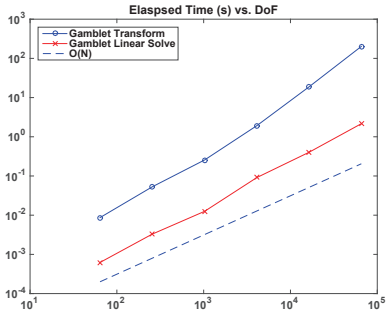


Figure 10: Elapsed time (sec) vs. DoF (N) for localized gamblet transform and localized gamblet linear solve, $\rho_k = 3$ for all k.

3 The wave PDE with rough coefficients

Consider the following prototypical wave PDE with rough coefficients,

$$\begin{cases} \mu(x)\partial_t^2 u(x,t) - \operatorname{div}\big(a(x)\nabla u(x,t)\big) = g(x,t) \quad x \in \Omega; \\ u(x,0) = u_0(x) \quad \text{on} \quad \Omega, \\ \partial_t u(x,0) = v_0(x) \quad \text{on} \quad \Omega, \\ u(x,t) = 0 \quad \text{on} \quad \partial\Omega \times [0,T] \end{cases} \tag{3.1}$$

where the domain Ω and the coefficients $\mu(x)$ and $a(x)$ are as in (2.1) (i.e. in $L^\infty(\Omega)$ and satisfy (2.2) and (2.3)).

Let $(\varphi_i)_{i\in\mathcal{N}}$ be a finite-dimensional (finite-element) basis of $H_0^1(\Omega)$ and write $\mathfrak{V} = \operatorname{span}\{\varphi_i \mid i \in \mathcal{N}\}$. Let $\tilde{u}(x,t) = \sum_{i\in\mathcal{N}} q_i(t)\varphi_i(x)$ be the finite-element solution of (3.1) in $\mathfrak{V}$ and assume that the elements $(\varphi_i)_{i\in\mathcal{N}}$ are chosen to satisfy (2.14) and (2.15) and so that $(\tilde{u}(x,t))_{0\le t\le T}$ is a *good enough* approximation of the solution $(u(x,t))_{0\le t\le T}$ of (3.1). Let $N = |\mathcal{N}|$ be the cardinal of $\mathcal{N}$ (and the dimension of $\mathfrak{V}$). Let M and K be the $\mathcal{N}\times\mathcal{N}$ mass and stiffness matrices $M_{i,j} = \int_\Omega \varphi_i\varphi_j\mu$ and $K_{i,j} = \int_\Omega (\nabla\varphi_i)^T a \nabla\varphi_j$.

Recall that the vector $q \in \mathbb{R}^{\mathbf{d}}$ is the solution of the forced Hamiltonian system

$$\begin{cases} \dot{q} &= M^{-1}p \\ \dot{p} &= -Kq + f \end{cases} \tag{3.2}$$

where for $i \in \mathcal{N}$, $f_i(t) := \int_\Omega \varphi_i g(x,t)$, $q_0 = q(0)$ corresponds to the coefficients of u_0 in the φ_i basis and $p_0 = p(0)$ is the $\mathcal{N}$-vector defined by $p_{0,i} := \int_\Omega \varphi_i v_0(x)\mu$.

3.1 Implicit midpoint rule

A popular time-discretization of (3.2) is the implicit midpoint rule [38], which is unconditionally stable (A-stable, i.e. its region of absolute stability includes the entire complex half-plane with negative real part), symplectic, symmetric (time-reversible) and preserves quadratic invariants exactly [38]. For example, when $f = 0$, (exactly preserved) quadratic invariants of (3.2) include the total energy ($E = \frac{1}{2}p^T M^{-1}p + \frac{1}{2}q^T Kq$) and the energy of each vibration mode ($E_i = |Q_i^T M^{-1}p|^2 \frac{1}{2}Q_i^T M Q_i + |Q_i^T q|^2 \frac{1}{2}Q_i^T K Q_i$ with $\lambda_i M Q_i = K Q_i$).

Writing q_n the numerical approximation of $q(n\Delta t)$, p_n the numerical approximation of $p(n\Delta t)$, and $f_n := f(n\Delta t)$, recall that the implicit midpoint

time discretization of (3.2) is

$$\begin{cases} q_{n+1} &= q_n + \Delta t M^{-1} \frac{p_n + p_{n+1}}{2} \\ p_{n+1} &= p_n - \Delta t K \frac{q_n + q_{n+1}}{2} + \Delta t f_{n+\frac{1}{2}} \end{cases} \tag{3.3}$$

Note that (3.3) can be written

$$\begin{cases} (M + \frac{(\Delta t)^2}{4} K) q_{n+1} &= (M - \frac{(\Delta t)^2}{4} K) q_n + \Delta t p_n + \Delta t^2 \frac{f_{n+\frac{1}{2}}}{2} \\ p_{n+1} &= p_n - \Delta t K \frac{q_n + q_{n+1}}{2} + \Delta t f_{n+\frac{1}{2}} \end{cases} \tag{3.4}$$

Let $u_n(x) := \sum_{i \in \mathcal{N}} q_{n,i} \varphi_i(x)$ and $v_n(x) := \sum_{i \in \mathcal{N}} (M^{-1} p_n)_i \varphi_i$ be the corresponding approximations of $u(x, n\Delta t)$ and $\partial_t u(x, n\Delta t)$. Observe that $p_{n,i} = \int_\Omega \varphi_i v_n \mu$ and solving (3.4) is equivalent to obtaining the finite element solution (in $\mathfrak{V}$) of

$$\begin{cases} \frac{4}{\Delta t^2} \mu u_{n+1} - \operatorname{div}\left(a \nabla u_{n+1}\right) &= \frac{4}{(\Delta t)^2} \mu u_n + \operatorname{div}\left(a \nabla u_n\right) + \frac{4}{\Delta t} \mu v_n + 2 g_{n+\frac{1}{2}} \\ \mu v_{n+1} &= \mu v_n^{\zeta,\mathrm{loc}} + \Delta t \operatorname{div}\left(a \nabla \frac{u_n + u_{n+1}}{2}\right) + \Delta t g_{n+\frac{1}{2}} \end{cases} \tag{3.5}$$

with $g_n(x) := g(x, n\Delta t)$.

3.2 Acceleration of the midpoint rule with ζ-gamblets

To achieve near linear complexity in the implementation of the midpoint rule we will perform the inversion of the implicit system in (3.4) or (3.5) in a localized ζ-gamblet basis with $\zeta = \Delta t$. Write $(q_n^{\mathrm{ap}}, p_n^{\mathrm{ap}})$ the output of the corresponding Algorithm 5.

Algorithm 5 Implicit midpoint rule with localized gamblets.

1: Set $\zeta = \Delta t$ and ε as in Theorem 3.1.
2: Compute $\chi_i^{(k),\zeta,\mathrm{loc}}, \psi_i^{(k),\zeta,\mathrm{loc}}, B^{(k),\zeta,\mathrm{loc}}, R^{(k,k-1),\zeta,\mathrm{loc}}$ with Algorithm 3.
3: $q_0^{\mathrm{ap}} := q_0$ and $p_0^{\mathrm{ap}} := p_0$.
4: **for** $n = 0$ to $T/\Delta t - 1$ **do**
5: Solve $(M + \frac{(\Delta t)^2}{4} K) q_{n+1}^{\mathrm{ap}} = (M - \frac{(\Delta t)^2}{4} K) q_n^{\mathrm{ap}} + \Delta t p_n^{\mathrm{ap}} + \frac{\Delta t^2}{2} f_{n+\frac{1}{2}}$ with Algorithm 4 // $f_{n+\frac{1}{2},i} := \int_\Omega \varphi_i g(x, (n+\frac{1}{2})\Delta t)$
6: $p_{n+1}^{\mathrm{ap}} = p_n^{\mathrm{ap}} - \Delta t K \frac{q_n^{\mathrm{ap}} + q_{n+1}^{\mathrm{ap}}}{2} + \Delta t f_{n+\frac{1}{2}}^{\mathrm{ap}}$.
7: **end for**

Write $u_n^{\rm ap} := \sum_{i\in\mathcal{N}} q_{n,i}^{\rm ap}\varphi_i$ and $v_n^{\rm ap} := \sum_{i\in\mathcal{N}}(M^{-1}p_n^{\rm ap})_i\varphi_i$. Observe that Line 5 of Algorithm 5 is equivalent to solving

$$\frac{4}{\Delta t^2}\mu u_{n+1}^{\rm ap} - \operatorname{div}\left(a\nabla u_{n+1}^{\rm ap}\right) = \frac{4}{(\Delta t)^2}\mu u_n^{\rm ap} + \operatorname{div}\left(a\nabla u_n^{\rm ap}\right) + \frac{4}{\Delta t}\mu v_n^{\rm ap} + 2g_{n+\frac{1}{2}} \tag{3.6}$$

in $\mathfrak{V}$ with Algorithm 4. Note that $u_{n+1}^{\rm ap} \in \mathfrak{V}$ and the equality (3.6) is defined in the finite-element sense after integration against $\varphi \in \mathfrak{V}$. Note also that (1) $p_{n,i}^{\rm ap} = \int_\Omega \varphi_i v_n^{\rm ap}\mu$, (2) $v_n^{\rm ap}$ is an approximation of $\partial_t u(x, n\Delta t)$, (3) $(q_n^{\rm ap})^T K q_n^{\rm ap} = \int_\Omega (\nabla u_n^{\rm ap})^T a \nabla u_n^{\rm ap}$, (4) $(p_n^{\rm ap})^T M^{-1} p_n^{\rm ap} = \int_\Omega (v_n^{\rm ap})^2\mu$.

The following theorem, whose proof is given in Subsection 7.1 of the appendix, provides a priori error estimates on the accuracy of Algorithm 5 using the exact midpoint rule solution as a reference. We assume for the clarity of those error estimates, without loss of generality, that $\Delta t \leq 1$ and write $\|g\|_{L^\infty(0,T,L^2(\Omega))}$) the essential supremum of $\|g(\cdot,t)\|_{L^2(\Omega)}$ over $t \in (0,T)$.

Theorem 3.1. *Let $u_n^{\rm ap}$ and $v_n^{\rm ap}$ be the output of Algorithm 5 and u_n, v_n be the solution of the implicit midpoint time discretization of* (3.5) *(or equivalently* (3.4)*) with time-step Δt. If ε in Algorithm 5 satisfies $\varepsilon \leq C^{-1}\frac{1}{T}\Delta t^3 h(\Delta t + h)$, then for $n\Delta t \leq T$ we have*

$$\begin{aligned}\|u_n - u_n^{\rm ap}\|_{H_0^1(\Omega)} + \|v_n - v_n^{\rm ap}\|_{L^2(\Omega)} \leq C(\Delta t)^2 \\ \left(\|u_0\|_{H_0^1(\Omega)} + \|v_0\|_{L^2(\Omega)} + T\|g\|_{L^\infty(0,T,L^2(\Omega))}\right)\end{aligned} \tag{3.7}$$

3.3 Numerical experiments

Let $a(x)$ be defined as in Example 2.2, $g(x,t) = \sin(2\pi(t+x_1))\cos(2\pi(t+x_2))$, $u(x,0) = 0$, and $u_t(x,0) = \sin(2\pi x_1)\cos(2\pi x_2)$. The reference solution is computed using bilinear finite-elements $\{\varphi_i \mid i \in \mathcal{N}\}$, and *Matlab* built-in integrator *ode15s* with time step $dt = 1/1280$. We test the performance/accuracy of exact gamblets and localized gamblets adapted to the implicit midpoint rule. We compute numerical solutions up to time $T = 1$. Figure 11 shows the reference solution, the numerical solution and the error of the numerical solution. The numerical solution is computed using localized gamblet for 2 stages Gauss-Legendre scheme (which will be introduced in Section § 5.1) with 2 layers, namely, we take $nl := \rho_k = 2$ in Algorithm 3 for $k = 1, \ldots, r$ and we will keep using the notation nl (number of layers) for the values of ρ_k, and $\Delta t = 0.1$.

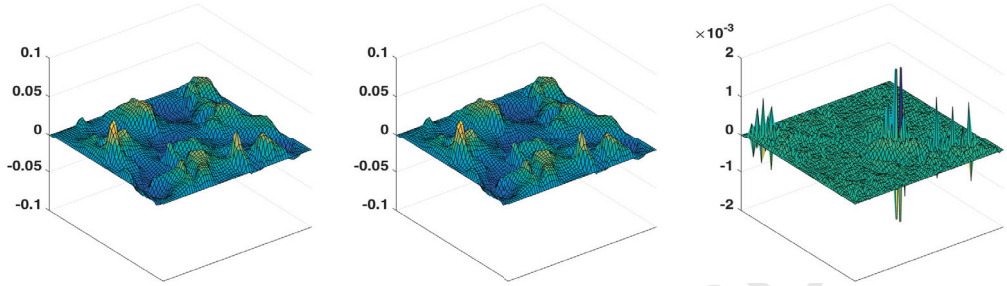


Figure 11: Solutions at $T = 1$. Left: Reference solution with the $\{\varphi_i \mid i \in \mathcal{N}\}$ basis, $dt = 1/1280$. Middle: numerical solution using localized gamblets with 3 layers for 2 stages Gauss-Legendre scheme, $\Delta t = 0.1$. Right: The error of numerical solution.

Figure 12 shows the relative error of the energy ($E = \frac{1}{2}p^T M^{-1} p + \frac{1}{2}q^T K q$) of gamblet solutions with respect to time and localization. The error for the gamblet solutions appears to be stable with respect to time for both the implicit midpoint scheme and the 2 stages Gauss-Legendre scheme if $nl > 1$. When $nl \geq 3$ for implicit midpoint and $nl \geq 4$ for 2 stages Gauss Legendre, the localized gamblet solutions are almost as accurate as the exact gamblet solutions.

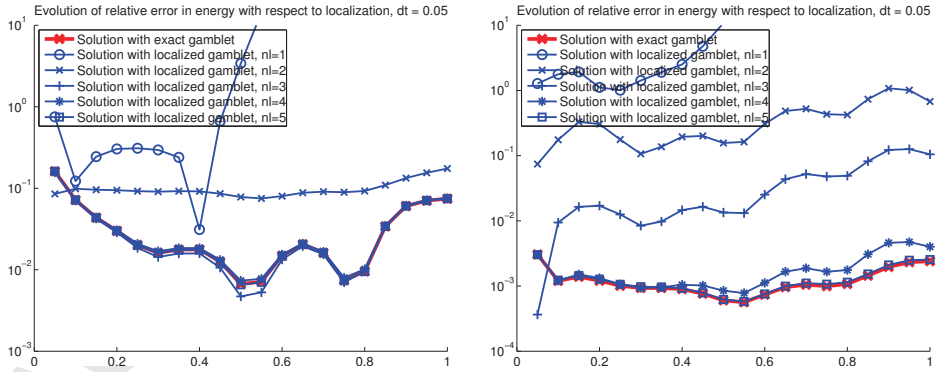


Figure 12: Evolution of the relative error of the energy w.r.t localization: Left, implicit midpoint scheme; Right, 2 stages Gauss-Legendre scheme.

Figure 13 and Figure 14 compare the components of the reference solution and the localized gamblet solution in each subband $\mathfrak{W}^{(k)}$, with implicit midpoint scheme and 2 stages Gauss-Legendre scheme respectively. A phase

shift error can be observed at high frequencies, and 2 stages Gauss-Legendre scheme has smaller phase error, even after localization.

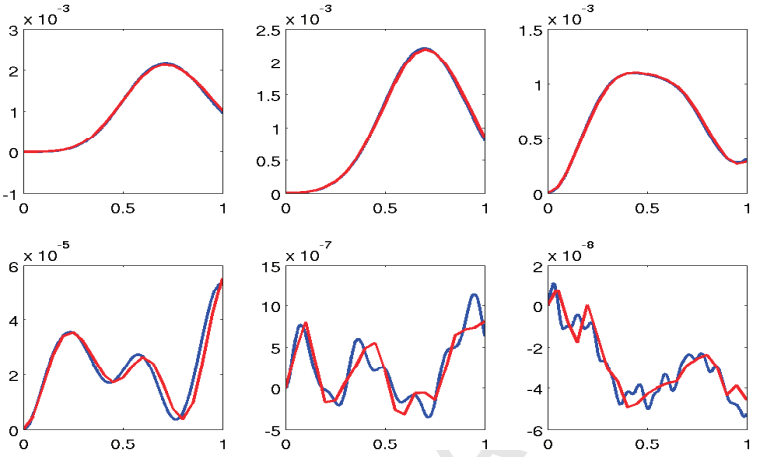


Figure 13: Evolution of $\chi_1^{(k)}$ component in subband $\mathfrak{W}^{(k)}$, $k = 1, \cdots, 6$, with localization parameter $nl = 3$, using implicit midpoint scheme. The blue curve is for the reference solution, and the red curve is for the localized gamblet solution.

Remark 3.2. *Figures 13 and 14 show that the multiresolution decomposition of the solution space performed by gamblets is analogous to a eigensubspace decomposition: the coefficients of the solution in $\mathfrak{W}^{(k),\zeta}$ evolve slowly for k small and fast for k large. Furthermore, these coefficients are robust to perturbations in initial conditions and dispersion errors for k small and sensitive for k large. Therefore gamblets decompose the the solution into components characterized by a hierarchy of levels of robustness. Furthermore, as done with wavelets [22], gamblet refinement could be used as an alternative to adaptive mesh refinement near singularities.*

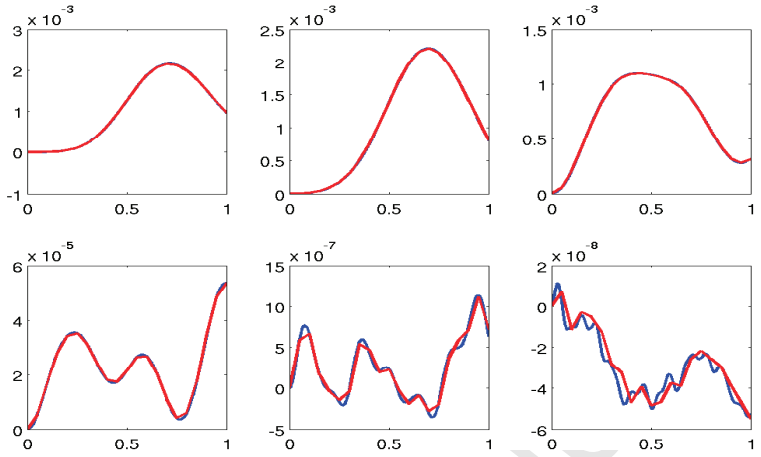


Figure 14: Evolution of $\chi_1^{(k)}$ component in subband $\mathfrak{W}^{(k)}$, $k = 1, \cdots, 6$, with localization parameter $nl = 3$, using 2 stages Gauss Legendre scheme. The blue curve is for the reference solution, and the red curve is for the localized gamblet solution.

Figure 15 shows the H^1 and L^2 errors at $T = 1$ with respect to localization for localized gamblet solutions, with time step $\Delta t = 0.025$. This shows that for fixed spatial resolution, the errors get saturated after reaching a critical nl. For example, we can choose $nl = 3$ as this critical value for the results shown in Figure 15 .

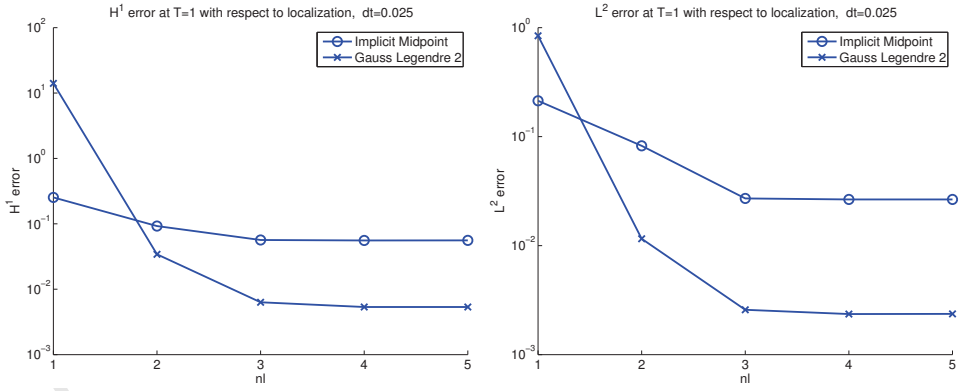


Figure 15: Error for localized gamblet solutions at $T = 1$: Left, H^1 error w.r.t localization; Right, L^2 error w.r.t. localization.

Figure 16 shows the H^1 and L^2 errors for the solution with exact gam-

blets and localized gamblets ($nl = 2$ or 3) using implicit midpoint scheme and 2 stages Gauss Legendre scheme at time $T = 1$, and time steps $\Delta t = 1/10, 1/20, 1/40, 1/80, 1/160, 1/320$.

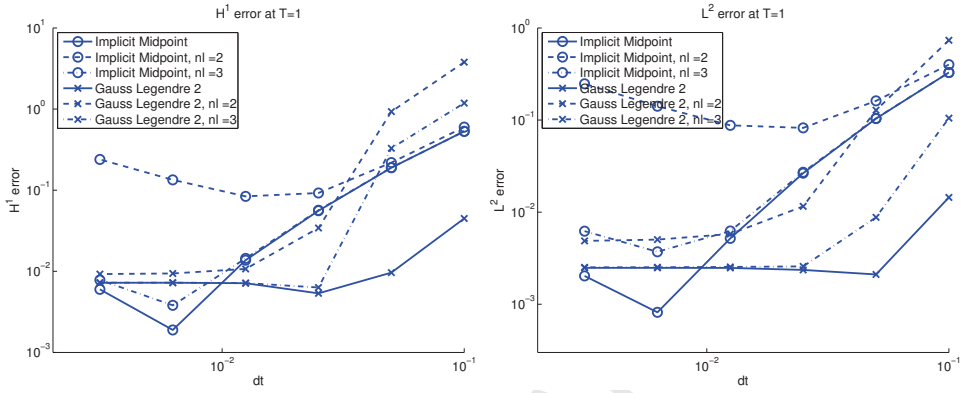


Figure 16: Left: H^1 error at time $T = 1$. Right: L^2 error at time $T = 1$.

The 2 stages Gauss Legendre scheme shows better performance for larger time steps and the localized gamblet solution is close to the exact gamblet solution for $nl = 3$.

4 The parabolic PDE with rough coefficients

Consider the following prototypical example of the parabolic PDE with rough coefficients

$$\begin{cases} \mu(x)\partial_t u(x,t) - \operatorname{div}\big(a(x)\nabla u(x,t)\big) = g(x,t) \quad x \in \Omega; \\ u(x,0) = u_0(x) \quad \text{on} \quad \Omega, \\ u(x,t) = 0 \quad \text{on} \quad \partial\Omega \times [0,T], \end{cases} \tag{4.1}$$

where the domain Ω and the coefficients $\mu(x)$ and $a(x)$ are as in (2.1) (i.e. in $L^\infty(\Omega)$ and satisfy (2.2) and (2.3)).

As in Section 3, we consider $(\varphi_i)_{i\in\mathcal{N}}$ a finite-dimensional (finite-element) basis of $H_0^1(\Omega)$, write $\tilde{u}(x,t) = \sum_{i\in\mathcal{N}} q_i(t)\varphi_i(x)$ the finite-element solution of (4.1) in $\mathfrak{V}$ (using the notations of Section 3) and assume that the elements $(\varphi_i)_{i\in\mathcal{N}}$ are chosen to satisfy (2.14) and (2.15) and so that $(\tilde{u}(x,t))_{0\le t\le T}$ is a *good enough* approximation of the solution of (4.1).

Recall that q is the solution of the ODE

$$M\dot{q} + Kq = f \tag{4.2}$$

where f is defined as in Section 3 and $q_0 = q(0)$ corresponds to the coefficients of u_0 in the φ_i basis.

4.1 Implicit-Euler time discretization

The implicit Euler time-discretization of (4.2) is

$$(M + \Delta t K) q_{n+1} = M q_n + \Delta t f_{n+1} \tag{4.3}$$

Recall that implicit Euler is first-order accurate and, in addition to being A-stable, is also L-stable and B-stable (see Definition 7.3 of Section 7) which are desirable for dissipative systems.

Let $u_n(x) := \sum_{i\in\mathcal{N}} q_{n,i}\varphi_i(x)$, be the corresponding approximation of $u(x, n\Delta t)$. Observe that solving (4.3) is equivalent to obtaining the finite element solution (in $\mathfrak{V}$) of

$$\frac{1}{\Delta t}\mu u_{n+1} - \operatorname{div}\left(a\nabla u_{n+1}\right) = \frac{1}{\Delta t}\mu u_n + g_{n+1} \tag{4.4}$$

with $g_n(x) := g(x, n\Delta t)$.

As in Subsection 3.2, to achieve near linear complexity in the implementation of the implicit Euler method we will perform the inversion of implicit system in (4.3) or (4.4) in a localized ζ-gamblet basis with $\zeta = 2\sqrt{\Delta t}$. Write $q_n^{\rm ap}$ the output of the corresponding Algorithm 6. Write

Algorithm 6 Implicit Euler with localized gamblets.

1: Set $\zeta = \Delta t$ and $\varepsilon = \Delta t^3$.
2: Compute $\chi_i^{(k),\zeta,\rm loc}, \psi_i^{(k),\zeta,\rm loc}, B^{(k),\zeta,\rm loc}, R^{(k,k-1),\zeta,\rm loc}$ with Algorithm 3.
3: $q_0^{\rm ap} := q_0$ // $u_0^{\rm ap} = u_0 = \sum_{i\in\mathcal{N}} q_{0,i}\varphi_i$.
4: **for** $n = 0$ to $T/\Delta t - 1$ **do**
5: Solve $(M + \Delta t K) q_{n+1}^{\rm ap} = M q_n^{\rm ap} + f_{n+1}$ with Algorithm 4 // $f_{n,i} := \int_\Omega \varphi_i g(x, n\Delta t)$
6: **end for**

$u_n^{\rm ap} := \sum_{i\in\mathcal{N}} q_{n,i}^{\rm ap}\varphi_i$. Observe that Line 5 of Algorithm 6 is equivalent to solving $\frac{1}{\Delta t}\mu u_{n+1}^{\rm ap} - \operatorname{div}\left(a\nabla u_{n+1}^{\rm ap}\right) = \frac{1}{\Delta t}\mu u_n^{\rm ap} + g_{n+1}$ in $\mathfrak{V}$ with Algorithm 4.

Theorem 4.1. *Let $u_n^{\rm ap}$ be the output of Algorithm 6, and u_n be the solution of implicit Euler time discretization of* (4.2) *with time-step Δt. It holds true that for $n \le T/\Delta t$,*

$$\|u_n^{\rm ap} - u_n\|_{H_0^1(\Omega)} \le C(\frac{T}{\Delta t})^2 \varepsilon \|g\|_{L^\infty(0,T,H^{-1}(\Omega))} \tag{4.5}$$

where ε is the localization parameter in Algorithm 6.

We refer to Subsection 7.3 for the proof of Theorem 7.3.

Remark 4.2. *The proof of Theorem 4.1 is based on Inequality* (7.16) *which is the B-stability condition of Definition 7.3. It is easy to show that this inequality is sufficient for validity of Theorem 4.1. Similar results to Theorem 4.1 holds true for B-stable methods and they can also be accelerated by the localized gamblets (e.g., the DIRK methods and SDIRK methods presented in Subsection 4.3).*

Remark 4.3. *By truncating the propagation of the solution at higher frequencies (large k) in the generalized gamblet decomposition one obtains a numerical homogenization of the wave or parabolic equations (as in [71, 69, 72]) with sub-linear complexity under sufficient regularity of initial conditions and source terms.*

4.2 TR-BDF2 time discretization

The TR-BDF2 [5] time-discretization of (4.2) is

$$\begin{cases} (M+\frac{\gamma\Delta t}{2}K)q_{n+\gamma}=(M-\frac{\gamma\Delta t}{2}K)q_n+\Delta t\frac{f_n+f_{n+\gamma}}{2} \\ (M+\frac{1-\gamma}{2-\gamma}\Delta tK)q_{n+1}=\frac{Mq_{n+\gamma}}{\gamma(2-\gamma)}-\frac{(1-\gamma)^2}{\gamma(2-\gamma)}Mq_n+\frac{1-\gamma}{2-\gamma}\Delta tf_{n+1} \end{cases} \tag{4.6}$$

Recall that TR-BDF2 is A-stable, L-stable but neither algebraically stable nor B-stable [29]. It is second order accurate and belongs to the category of diagonally implicit Runge-Kutta (DIRK) methods. We select $\gamma=2-\sqrt{2}$ to minimize the local error [5] and ensure $\dfrac{\gamma}{2}=\dfrac{1-\gamma}{2-\gamma}$ (under that choice, (4.6) requires solving two systems of the same form $(M+\frac{\gamma\Delta t}{2}K)Q=b$ at each time step). Let $u_n(x):=\sum_{i\in\mathcal{N}}q_{n,i}\varphi_i(x)$ be the corresponding approximation of $u(x,n\Delta t)$. Observe that solving (4.6) is equivalent to obtaining the finite element solution (in $\mathfrak{V}$) of

$$\begin{cases} \frac{2}{\gamma\Delta t}\mu u_{n+\gamma}-\operatorname{div}\left(a\nabla u_{n+\gamma}\right)=\frac{2\mu u_n}{\gamma\Delta t}+\operatorname{div}\left(a\nabla u_n\right)+\frac{g_n+g_{n+\gamma}}{\gamma} \\ \frac{2}{\gamma\Delta t}\mu u_{n+1}-\operatorname{div}\left(a\nabla u_{n+1}\right)=\frac{1}{\gamma(1-\gamma)\Delta t}\mu u_{n+\gamma}-\frac{1-\gamma}{\gamma\Delta t}\mu u_n+g_{n+1} \end{cases} \tag{4.7}$$

As in Subsection 3.2, to achieve near linear complexity we will perform the inversion of implicit systems in (4.7) in a localized ζ-gamblet basis with $\zeta=\sqrt{2\gamma\Delta t}$.

4.3 DIRK3 and SDIRK3

Other popular time-discretion methods for (4.2) are DIRK3 [11] and SDIRK3 [16, p262]. DIRK3 (3-stages Diagonally Implicit Runge-Kutta) is L-stable and B-stable [11], and its Butcher tableau is given in Table 1. The implemen-

0.0585104413419415	0.0585104413426586	0.0	0.0
0.8064574322792799	0.0389225469556698	0.7675348853239251	0.0
0.2834542075672883	0.1613387070350185	-0.5944302919004032	0.7165457925008468
	0.1008717264855379	0.4574278841698629	0.4417003893445992

Table 1: Butcher tableau for DIRK3

tation of DIRK3 requires solving 3 equations $\frac{1}{\Delta t A_{i,i}}\mu w_i - \mathrm{div}\left(a\nabla w_i\right) = b_i$ using finite-elements in $\mathfrak{V}$, where $A_{1,1}, \ldots, A_{3,3}$ are the diagonal entries of the Runge-Kutta matrix A of DIRK3.

SDIRK3 (3-stage Singly Diagonally Implicit Runge-Kutta) is L-stable [16, P 262] and its Butcher tableau is given in Table 2 which has identical diagonal entries. The implementation of SDIRK3 requires solving 3 equations $\frac{1}{\Delta t\lambda}\mu w_i - \mathrm{div}\left(a\nabla w_i\right) = b_i$ using finite-elements in $\mathfrak{V}$, where λ is defined in Table 2. As in Subsection 3.2, to achieve near linear complexity

λ	λ	0	0
$\frac{1}{2}(1+\lambda)$	$\frac{1}{2}(1-\lambda)$	λ	0
1	$\frac{1}{4}(-6\lambda^2+16\lambda-1)$	$\frac{1}{4}(6\lambda^2-20\lambda+5)$	λ
	$\frac{1}{4}(-6\lambda^2+16\lambda-1)$	$\frac{1}{4}(6\lambda^2-20\lambda+5)$	λ

Table 2: Butcher tableau for SDIRK3 where $\lambda \simeq 0.4358665215$ (identified as a root of $\frac{1}{6} - \frac{3}{2}\lambda + 3\lambda^2 - \lambda^3 = 0$) ensures L-stability.

we will perform the inversion of linear systems of DIRK3 using 3 localized ζ-gamblets with $\zeta = \sqrt{A_{i,i}\Delta t/2}$, and the inversion of linear systems of SDIRK3 using only 1 localized ζ-gamblets with $\zeta = \sqrt{\lambda\Delta t/2}$.

4.4 Numerical experiments

Let $a(x)$ be defined as in Example 2.2 and $r = 6$, $g(x,t) = \sin(2\pi(t + x_1))\cos(2\pi(t+x_2))$, $u(x,0) = \sin(2\pi x_1)\cos(2\pi x_2)$. The reference solution is the finite element solution with piecewise bilinear elements $\{\varphi_i \mid i \in \mathcal{N}\}$, and *Matlab* built-in integrator *ode15s* with time step $dt = 1/1280$. We test the

performance of exact and localized gamblets adapted to implicit Euler, TR-BDF2, DIRK3, SDIRK3, and fully implicit Runge-Kutta methods Radau IIA and Lobatto IIIC (which will be introduced in § 5.2). We compute numerical solutions up to time $T = 1$. Figure 17 shows the reference solution, the DIRK3 solution with localized gamblet ($nl = 3$) and its numerical error with respect to the reference solution.

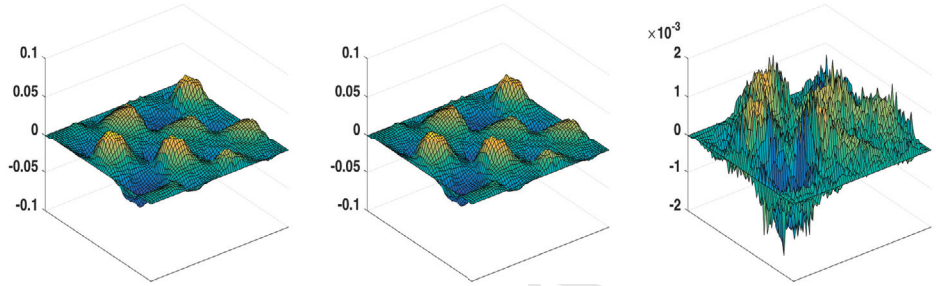


Figure 17: Solutions at $T = 1$. Left: Reference solution; Middle: numerical solution with localized gamblet, DIRK3, $nl = 3$ and $\Delta t = 0.05$; Right: error of the localized solution.

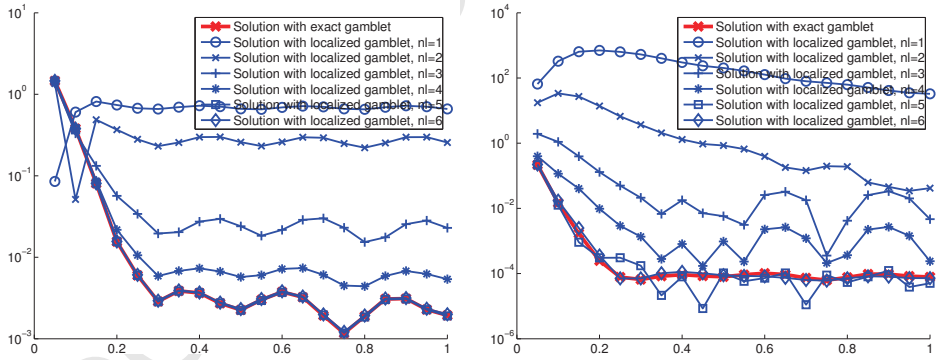


Figure 18: Evolution of the relative error of energy w.r.t localization ($\Delta t = 0.05$): Left, TR-BDF2; Right, Radau IIA scheme.

Figure 18 shows the relative error of the energy of gamblet solutions with respect to time and localization. The errors decays when the number of layers increases, and Radau IIA scheme has better accuracy compared to TR-BDF2 method, note that Radau IIA is a 5th order method and TR-BDF2 is a 2nd order method.

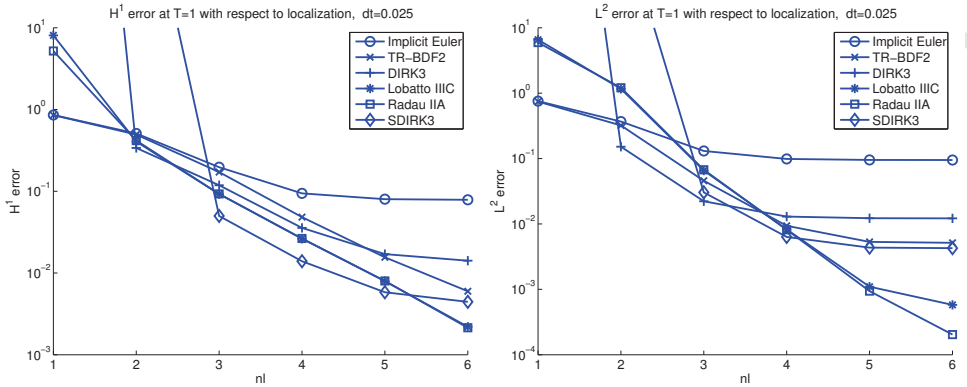


Figure 19: Left: H^1 error w.r.t localization parameter nl; Right: L^2 error w.r.t localization parameter nl. $\Delta t = 0.025$

Figure 19 shows the H^1 and L^2 errors at $T = 1$ for all the 6 numerical schemes with respect to different localization parameters $nl = 1, \cdots, 6$ for exact and localized gamblet solutions. High order methods such as Radau IIA and Lobatto IIIC have best accuracy if more localizaton layers are used. When $nl = 2$ or 3, it appears that simpler methods such as TR-BDF2, DIRK3 or SDIRK3 achieve a better balance between accuracy and computational cost.

Figure 20 compares the components of the reference solution and localized gamblet solutions in each subband $\mathfrak{W}^{(k),\zeta}$, computed with the DIRK3 scheme (we observe similar results for other schemes). Most of the error occurs in the first subband and at early time, and gets damped quickly.

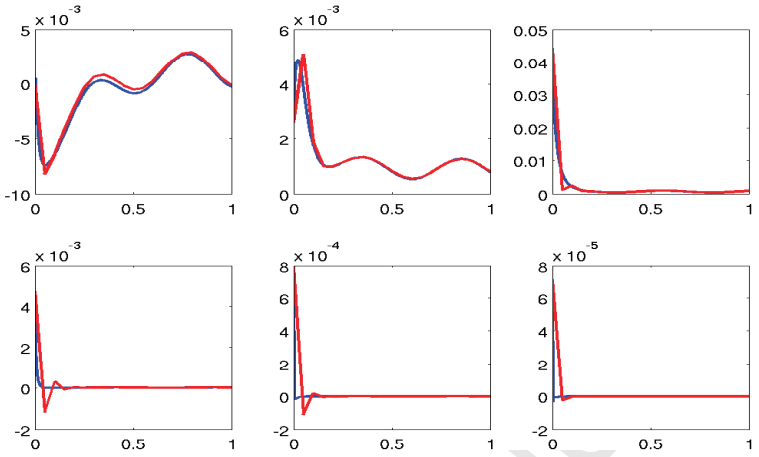


Figure 20: Evolution of $\chi_1^{(k)}$ component in subband $\mathfrak{W}^{(k),\zeta}$, $k = 1, \cdots, 6$, with localized gamblet ($nl = 3$) for DIRK3 scheme. The blue curve is obtained from the reference solution, and the red curve is obtained from the localized gamblet solution.

Figure 21 shows H^1 and L^2 errors with exact gamblets and localized gamblets ($nl = 3$) at time $T = 1$ with time steps $\Delta t = 1/10, 1/20, 1/40, 1/80, 1/160, 1/320$. All 6 methods are tested: implicit Euler, TR-BDF2, DIRK3, Lobatto IIIC, Radau IIA and SDIRK3. In general, the higher order methods such as Radau IIA and Lobatto IIIC are more accurate for coarser time steps, refinement of time steps does not reduce the error further due to the fixed spatial resolution. Localized gamblet solutions converge as time steps decrease. Localized TR-BDF2 and Implicit Euler have a slower convergence rate compared to higher order methods.

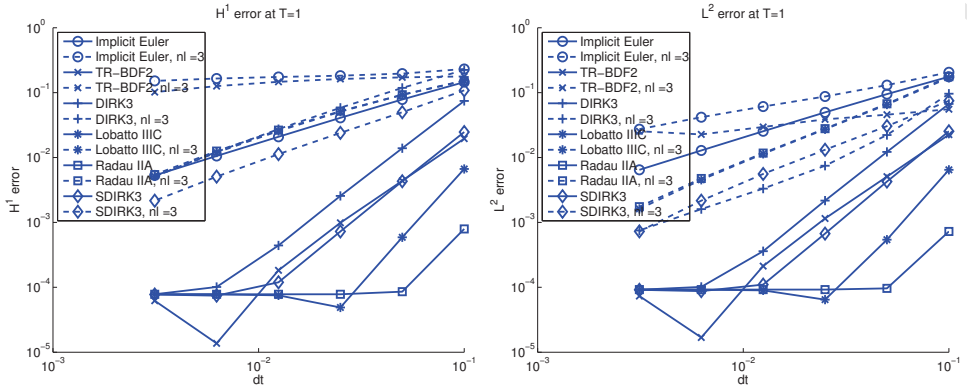


Figure 21: Left: H^1 error at time $T = 1$; Right: L^2 error at time $T = 1$.

5 Complex gamblets for higher order implicit schemes

5.1 Solving wave equation with 2 stages Gauss-Legendre scheme

The implicit midpoint scheme introduced in section § 3.1 is a 1 stage Gauss-Legendre method. To obtain higher order method for the wave equation, we can use higher order Gauss-Legendre methods . Here we describe the implementation of 2 stages Gauss-Legendre scheme, which is 4th order accurate, unconditionally stable, symplectic, symmetric (time-reversible) and preserves quadratic invariants exactly [38]. In particular, we will show how to use gamblets to achieve near linear complexity.

The Butcher tableau for the 2 stages Gauss-Legendre scheme is as follows. Namely, the Runge-Kutta matrix, weights and nodes are $A = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \end{pmatrix}$,

$$\begin{array}{c|cc} \frac{1}{2} - \frac{\sqrt{3}}{6} & \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{2} + \frac{\sqrt{3}}{6} & \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

Table 3: Butcher tableau for GL2

$b = (\frac{1}{2}, \frac{1}{2})$, and $c = (\frac{1}{2} - \frac{\sqrt{3}}{6}, \frac{1}{2} + \frac{\sqrt{3}}{6})^T$.

Using notations in § 3, let $y_n = (q_n; p_n)$ be a column vector of length $2N$, $f_n^1 = (\mathbf{0}; f(t_n + c_1 h))$, and $f_n^2 = (\mathbf{0}; f(t_n + c_2 h))$, where $\mathbf{0}$ is the zero valued

column vector of length N. Apply the 2 stages Gauss-Legendre scheme to equation (3.2), we have

$$\begin{aligned} k_n^1 &= f_n^1 + \begin{pmatrix} \mathbf{0} & M^{-1} \\ -K & \mathbf{0} \end{pmatrix} (y_n + \Delta t A_{11} k_n^1 + \Delta t A_{12} k_n^2) \\ k_n^2 &= f_n^2 + \begin{pmatrix} \mathbf{0} & M^{-1} \\ -K & \mathbf{0} \end{pmatrix} (y_n + \Delta t A_{21} k_n^1 + \Delta t A_{22} k_n^2) \\ y_{n+1} &= y_n + \Delta t (b_1 k_n^1 + b_2 k_n^2) \end{aligned}$$

Define H in the following tensor product format,

$$H := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} M & 0 \\ 0 & \mathbf{I} \end{pmatrix} + a \otimes \begin{pmatrix} 0 & -h\mathbf{I} \\ hK & 0 \end{pmatrix} \tag{5.1}$$

where $\mathbf{I}$ is the identity matrix of size N.

Write $F_n := (p_n; -Kq_n + f(t_n + c_1\Delta t); p_n; -Kq_n + f(t_n + c_1\Delta t))^T$, we need to solve the coupled linear system $Hk_n = F_n$ for $k_n = (k_n^1; k_n^2)$, and then y_{n+1}.

Since A is diagonalizable, we can write $A = S\Lambda S^{-1}$ with $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$. Use $T := S \otimes \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$ to block diagonalize H. That is,

$$\tilde{H} := THT^{-1} = \begin{pmatrix} M & -\Delta t\lambda_1 \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \Delta t\lambda_1 K & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & M & -\Delta t\lambda_2 \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \Delta t\lambda_2 K & \mathbf{I} \end{pmatrix} \tag{5.2}$$

where $T^{-1} = S^{-1} \otimes \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$.

Write $\tilde{F}_n := T^{-1}F_n$ and $\tilde{k}_n := T^{-1}k_n$, we have $\tilde{H}\tilde{k}_n = \tilde{F}_n$. Therefore, instead of solving the coupled linear system with respect to H, we solve the decoupled linear systems with respect to $\tilde{H}$. Let $\bar{H}_1 := \begin{pmatrix} M & -\Delta t\lambda_1 \mathbf{I} \\ \Delta t\lambda_1 K & \mathbf{I} \end{pmatrix}$ and $\bar{H}_2 := \begin{pmatrix} M & -\Delta t\lambda_2 \mathbf{I} \\ \Delta t\lambda_2 K & \mathbf{I} \end{pmatrix}$. Similar to the gamblet solution for implicit midpoint scheme (3.4), we only need to introduce gamblets associated with the matrices $\Delta t^2 \lambda_i^2 K + M$ to solve the linear systems associated with $\bar{H}_i$, $i = 1, 2$.

5.2 Solving parabolic equation with fully implicit Runge-Kutta methods

Let q be the solution of the semidiscrete ODE system derived from the parabolic PDE (4.1),

$$M\dot{q} + Kq = f \tag{5.3}$$

where f is defined as in Section § 3 and $q_0 = q(0)$ corresponds to the coefficients of u_0 in the φ_i basis. Higher order implicit Runge-Kutta methods can be used to solve (5.3) to achieve better stability and accuracy. Write A, b and c the Runge-Kutta matrix, weights, and nodes. Let s be the number of stages of the Runge-Kutta method. The Runge-Kutta method can be written as,

$$q_{n+1} = q_n + \Delta t \sum_{i=1}^{s} b_i k_n^i,$$

with

$$Mk_n^i = -K(q_n + \Delta t \sum_{i=1}^{s} A_{ij} k_n^j) + f(t_n + c_i \Delta t), \quad k = 1, \ldots, s$$

Write

$$H := \mathbf{I}_s \otimes M + \Delta t A \otimes K\,.$$

where $\mathbf{I}_s$ is the identity matrix of size s.

To obtain q_{n+1}, we need to solve the coupled linear system $Hk_n = F_n$ for $k_n := (k_n^1; \cdots; k_n^s)$, $F_n := (-Kq_n + Mf(t_n + c_1\Delta t); \ldots; -Kq_n + Mf(t_n + c_s\Delta t))$.

Assume A is diagonalizable, such that $A = S\Lambda S^{-1}$ with $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_s)$. Write $T := S \otimes \mathbf{I}$ (and $T^{-1} = S^{-1} \otimes \mathbf{I}$), we have,

$$\tilde{H} := THT^{-1} = \mathbf{I}_s \otimes M + \Delta t \Lambda \otimes K$$

Write $\tilde{k}_n := Tk_n$ and $\tilde{F}_n := TF_n$, we can use gamblets associated with the matrices $M + \Delta t \lambda_i K (i = 1, \ldots, s)$ to solve the decoupled linear system $\tilde{H}\tilde{k}_n = \tilde{F}_n$ for $\tilde{k}_n$, then obtain k_n and q_{n+1}.

In our numerical illustrations in § 4 we have considered the following fully implicit Runge-Kutta methods: Lobatto IIIC and Radau IIA. Recall that Lobatto IIIC [39, 16] is 4th order accurate, L-stable, B-stable, stiffly accurate, and its Butcher tableau is in Table 4.

Recall also that Radau IIA [39, 16] is 5th order accurate, A-stable and its Butcher tableau is in Table 5.

0	1/6	-1/3	1/6
0.5	1/6	5/12	-1/12
1	1/6	2/3	1/6
	1/6	2/3	1/6

Table 4: Butcher tableau for Lobatto IIIC scheme.

$\frac{2}{5} - \frac{\sqrt{6}}{10}$	$\frac{11}{45} - \frac{7\sqrt{6}}{360}$	$\frac{37}{225} - \frac{169\sqrt{6}}{1800}$	$-\frac{2}{225} + \frac{\sqrt{6}}{75}$
$\frac{2}{5} + \frac{\sqrt{6}}{10}$	$\frac{37}{225} + \frac{169\sqrt{6}}{1800}$	$\frac{11}{45} + \frac{7\sqrt{6}}{360}$	$-\frac{2}{225} - \frac{\sqrt{6}}{75}$
1	$\frac{4}{9} - \frac{\sqrt{6}}{36}$	$\frac{4}{9} + \frac{\sqrt{6}}{36}$	$\frac{1}{9}$
	$\frac{4}{9} - \frac{\sqrt{6}}{36}$	$\frac{4}{9} + \frac{\sqrt{6}}{36}$	$\frac{1}{9}$

Table 5: Butcher tableau for Radau IIA scheme.

5.3 Gamblet transformation for complex valued matrix

As shown in § 5.1 and § 5.2, we need to introduce generalized gamblets for matrices of the form $M + \Delta t^2 \lambda^2 K$ for hyperbolic equations and $M + \Delta t \lambda K$ for parabolic equations to open the complexity bottleneck of higher order implicit schemes, where λ is the eigenvalue of the corresponding Runge-Kutta matrix. For implicit Runge-Kutta methods, such as Gauss-Legendre type, Lobatto type, and Radau type, λ is in general a complex number.

Definition 5.1. *The definition of complex valued ζ-gamblets is algebraically identical to that of real-value ζ-gamblets. We keep using the scalar product defined in* (2.7)*, and the notion of orthogonality remains the same as in* (2.7)*. Algorithm 1 and 3 remain unchanged, in particular we do not replace matrix transpose operations with complex conjugate transpose operations.*

Remark 5.2. *It is a simple observation that the gamblet transform remains algebraically exact for complex valued matrices (this can be observed rigorously and numerically). Although, we loose the positivity of the scalar product* (2.7) *when ζ is complex (the matrices $M + \Delta t^2 \lambda^2 K$ and $M + \Delta t \lambda K$ remain symmetric when λ is a complex complex number but they are neither positive definite nor Hermitian), figures 24 and 25, show that, for Radau IIA (complex ζ) and SDIRK3 (real ζ), condition numbers and ranges of eigenvalues remain similar. Although this is not proven in this paper, we suspect that the preservation of uniformly bounded condition numbers and exponential decay with complex valued ζ is generic.*

Figures 22 and 23 illustrate the real part and imaginary parts of the

gamblet basis associated with the first complex eigenvalue (0.1626+0.1849i) of Radau IIA and $\Delta t = .1$.

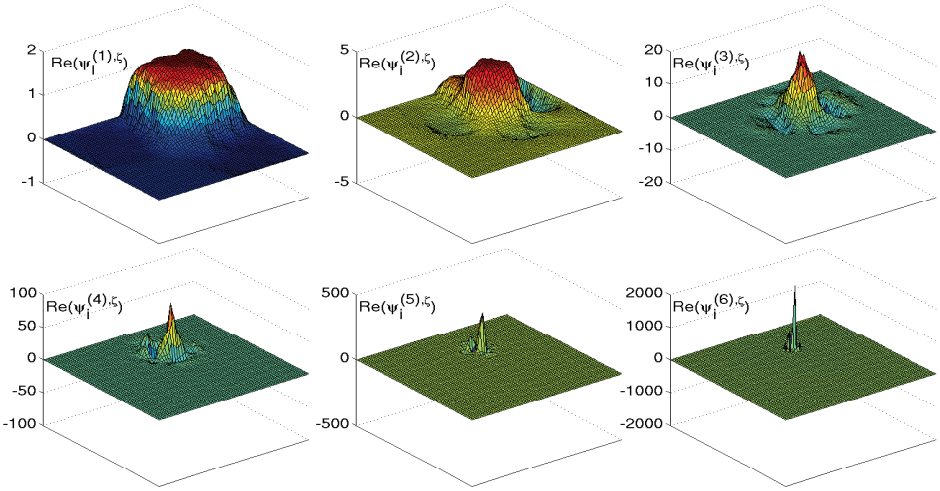


Figure 22: Real part of χ_i^k associated with the first complex eigenvalue of Radau IIA and $\Delta t = .1$ for parabolic equation.

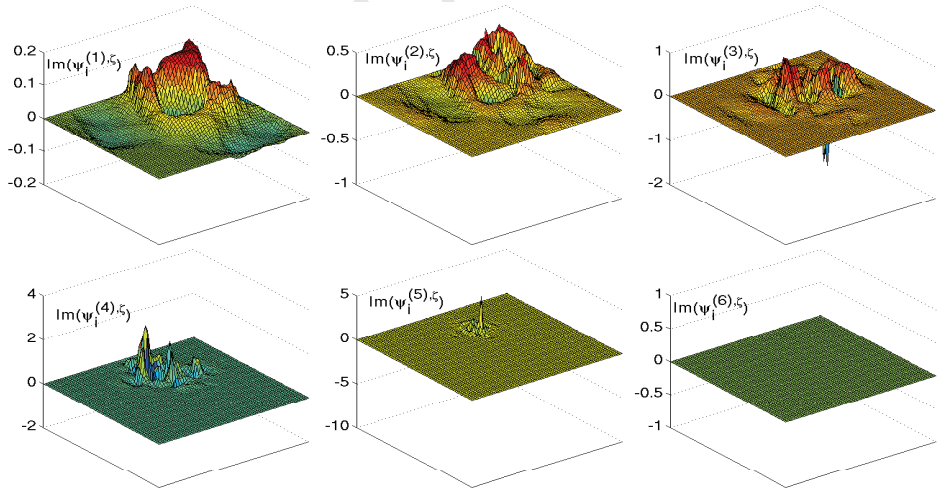


Figure 23: Imaginary part of χ_i^k associated with the first complex eigenvalue of Radau IIA and $\Delta t = .1$ for parabolic equation.

Figures 24 and 25 compare the condition numbers and the ranges of eigenvalues in $\mathfrak{W}^{(k)}$ of the complex gamblets associated with the first com-

plex eigenvalue (0.1626+0.1849i) of the RK matrix of Radau IIA and the real gamblets associated with the first eigenvalue (0.4359) of the RK matrix of SDIRK3. Note the similarity between the condition numbers and ranges of eigenvalues of the complex gamblets of Radau IIA and the real gamblets of SDIRK3.

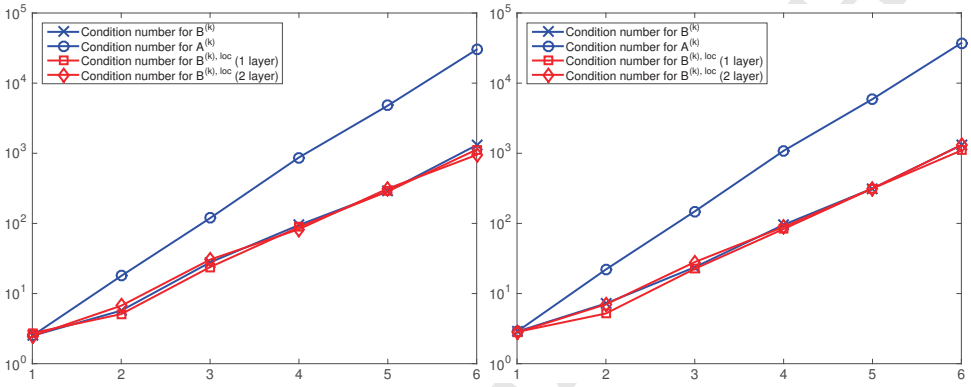


Figure 24: Left: Range of (the modulus of complex) eigenvalues in $\mathfrak{W}^k$ associated with the first eigenvalue of Radau IIA and $\Delta t = 1$. Right: Range of (real) eigenvalues in $\mathfrak{W}^k$ associated with SDIRK3 and $\Delta t = 1$.

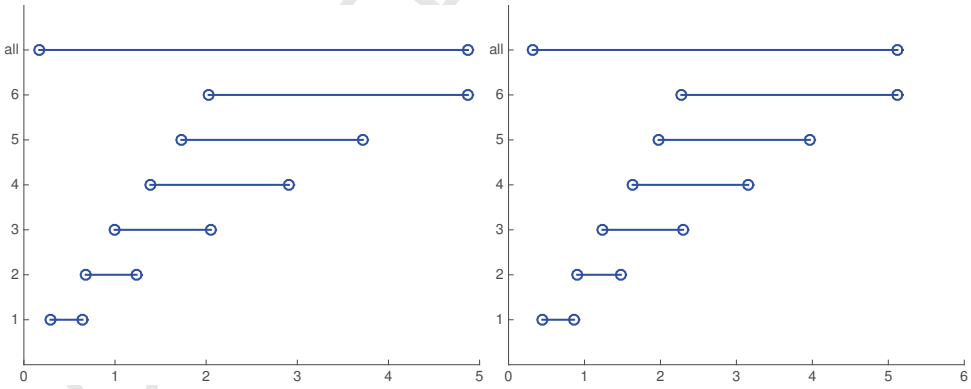


Figure 25: Left: Range of (the modulus of complex) eigenvalues in $\mathfrak{W}^k$ associated with the first eigenvalue of Radau IIA and $\Delta t = 1$. Right: Range of (real) eigenvalues in $\mathfrak{W}^k$ associated with SDIRK3 and $\Delta t = 1$.

6 Solving the parabolic equation in near-linear complexity with multi-time-stepping

As illustrated in Figure 20, errors in finer subbands (corresponding to larger eigenvalues) decay quickly. Therefore by refining time-steps close to the final stop time it is possible to lower the computational complexity and still preserve the accuracy. Motivated by this observation we propose the following $\mathcal{O}(N \ln^{3d+1} N)$ complexity multi-time-stepping algorithm for solving (5.3) (up to grid-size accuracy in energy norm). Let $T = M\Delta t$, prescribe an error threshold $\varepsilon < \Delta t$, then there exist $s \in \mathbb{N}$, such that $\frac{\Delta t}{2^s} \leq \varepsilon < \frac{\Delta t}{2^{s-1}}$. For $n = 1, 2, ..., M-1$, we use gamblets associated with time step Δt to obtain q_n up to $T - \Delta t$. For the last (coarse) time step from $T - \Delta t$ to T, we subsequently choose time steps $\Delta t/2$, $\Delta t/4$, ..., $\Delta t/2^s$, $\Delta t/2^s$, and solve the implicit scheme with gamblets associated to those time steps.

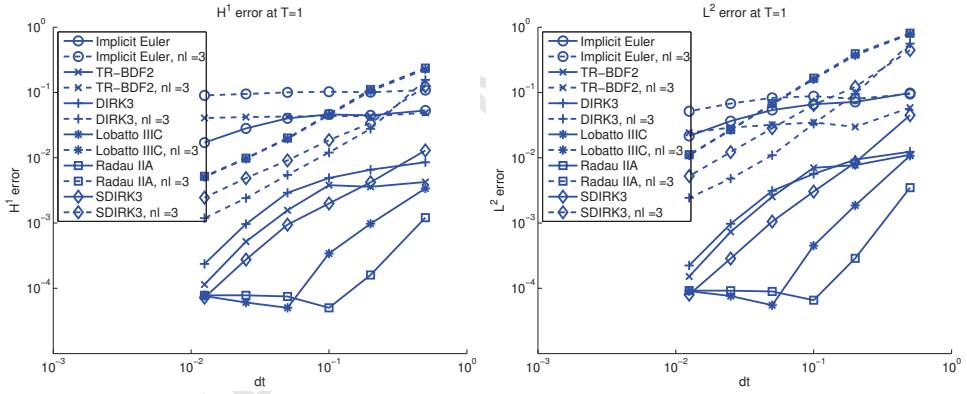


Figure 26: Solving parabolic equation with multi-time-stepping, $\varepsilon = 1/1280$, Left: H^1 error at time $T = 1$; Right: L^2 error at time $T = 1$.

Comparing Figures 26 and 21, localized gamblets (with 3 layers) achieve H^1 error of 10^{-3} with multi-time-stepping and $\Delta t \simeq 0.02$, we need $\Delta t \simeq 0.004$ to achieve the same accuracy with uniform time stepping.

7 Appendix

7.1 Proof of Theorem 3.1

We will need the following lemma,

Lemma 7.1. *Let (q_n, p_n) be the solution of (3.3). Write $E_n := \frac{1}{2}p_n^T M^{-1} p_n + \frac{1}{2}q_n^T K q_n$. Using the notation $|f|_{M^{-1}} := \sqrt{f^T M^{-1} f}$ we have*

$$|\sqrt{E_n} - \sqrt{E_0}| \leq \Delta t 2^{-1/2} \sum_{k=0}^{n-1} |f_{k+\frac{1}{2}}|_{M^{-1}} \tag{7.1}$$

Proof. Multiplying the first line of (3.3) by $(q_{n+1} + q_n)^T K$, the second line by $(p_{n+1} + p_n)^T M^{-1}$ and summing together, we obtain that

$$E_{n+1} - E_n = \Delta t (p_{n+1} + p_n)^T M^{-1} \frac{f_{n+\frac{1}{2}}}{2}.$$

Observe that

$$|(p_{n+1}+p_n)^T M^{-1} f_{n+\frac{1}{2}}| \leq |p_{n+1}+p_n|_{M^{-1}} |f_{n+\frac{1}{2}}|_{M^{-1}} \leq \sqrt{2}(\sqrt{E_{n+1}}+\sqrt{E_n})|f_{n+\frac{1}{2}}|_{M^{-1}}.$$

We have $|\sqrt{E_{n+1}} - \sqrt{E_n}| \leq \Delta t 2^{-1/2} |f_{n+\frac{1}{2}}|_{M^{-1}}$, and we conclude the proof by induction. □

Let us now prove Theorem 3.1. Line 5 of Algorithm 5 and Theorems 2.17 and 2.18 imply that $(\frac{4}{(\Delta t)^2}M + K)q_{n+1}^{\mathrm{ap}} = b_n + (\frac{4}{(\Delta t)^2}M + K)e_n$ with $b_n = (\frac{4}{(\Delta t)^2}M - K)q_n^{\mathrm{ap}} + \frac{4}{\Delta t}p_n^{\mathrm{ap}} + 2f_{n+\frac{1}{2}}$ and

$$e_n^T (\frac{4}{(\Delta t)^2}M + K) e_n \leq C\varepsilon^2 b_n^T K^{-1} b_n \tag{7.2}$$

Therefore, lines 5 and 6 of Algorithm 5 can be written as,

$$\begin{cases} q_{n+1}^{\mathrm{ap}} - q_n^{\mathrm{ap}} &= \Delta t M^{-1} \frac{p_n^{\mathrm{ap}} + p_{n+1}^{\mathrm{ap}}}{2} + s_n \\ p_{n+1}^{\mathrm{ap}} - p_n^{\mathrm{ap}} &= -\Delta t K \frac{q_n^{\mathrm{ap}} + q_{n+1}^{\mathrm{ap}}}{2} + \Delta t f_{n+\frac{1}{2}} \end{cases} \tag{7.3}$$

with

$$s_n = \frac{(\Delta t)^2}{4} M^{-1} (\frac{4}{(\Delta t)^2}M + K) e_n \tag{7.4}$$

Write $q_n^{\mathrm{err}} := q_n^{\mathrm{ap}} - q_n$ and $p_n^{\mathrm{err}} := p_n^{\mathrm{ap}} - p_n$, together with (7.3) with (7.4), it leads to

$$\begin{cases} q_{n+1}^{\mathrm{err}} - q_n^{\mathrm{err}} &= \Delta t M^{-1} \frac{p_n^{\mathrm{err}} + p_{n+1}^{\mathrm{err}}}{2} + s_n \\ p_{n+1}^{\mathrm{err}} - p_n^{\mathrm{err}} &= -\Delta t K \frac{q_n^{\mathrm{err}} + q_{n+1}^{\mathrm{err}}}{2} \end{cases} \tag{7.5}$$

Write $E_n^{\rm err} := \frac{1}{2} p_n^{{\rm err},T} M^{-1} p_n^{\rm err} + \frac{1}{2} q_n^{{\rm err},T} K q_n^{\rm err}$. Left multiplying the first equation of (7.5) by $\frac{1}{2}(q_{n+1}^{\rm err} + q_n^{\rm err})^T K$ and the second equation by $\frac{1}{2}(p_{n+1}^{\rm err} + p_n^{\rm err}) M^{-1}$, then adding the resulting equations we obtain that

$$E_{n+1}^{\rm err} - E_n^{\rm err} = \frac{1}{2}(q_{n+1}^{\rm err} + q_n^{\rm err})^T K s_n \tag{7.6}$$

Observing that $|(q_{n+1}^{\rm err} + q_n^{\rm err})^T K s_n| \le \sqrt{2}(\sqrt{E_{n+1}^{\rm err}} + \sqrt{E_n^{\rm err}})|s_n|_K$, we have

$$|\sqrt{E_{n+1}^{\rm err}} - \sqrt{E_n^{\rm err}}| \le 2^{-\frac{1}{2}} |s_n|_K \tag{7.7}$$

Write $B := \frac{4}{(\Delta t)^2} M + K$, (7.4) and (7.2) imply that

$$(\frac{4}{(\Delta t)^2})^2 s_n^T M B^{-1} M s_n \le C\varepsilon^2 b_n^T K^{-1} b_n \,. \tag{7.8}$$

We also have

$$s_n^T K s_n \le \frac{\lambda_{\max}(K)\lambda_{\max}(B)}{(\lambda_{\min}(M))^2} s_n^T M B^{-1} M s_n, \tag{7.9}$$

and

$$\begin{aligned} |b_n|_{K^{-1}} \le & \big(\frac{4}{(\Delta t)^2}\frac{\lambda_{\max}(M)}{\lambda_{\min}(K)} + 1\big)|q_n^{\rm ap}|_K + \frac{4}{\Delta t}\sqrt{\frac{\lambda_{\max}(M)}{\lambda_{\min}(K)}}|p_n^{\rm ap}|_{M^{-1}} \\ & + 2\sqrt{\frac{\lambda_{\max}(M)}{\lambda_{\min}(K)}}\|g_{n+\frac{1}{2}}\|_{L^2(\Omega)}. \end{aligned} \tag{7.10}$$

Poincaré's inequality, (2.14) and (2.15) lead to

$$\lambda_{\max}(M) \le CN^{-1}, \qquad C^{-1}N^{-1} \le \lambda_{\min}(M), \tag{7.11}$$

$$\lambda_{\max}(K) \le CN^{-1}h^{-2}, \qquad C^{-1}N^{-1} \le \lambda_{\min}(K)\,. \tag{7.12}$$

Summarizing we have obtained that $|s_n|_K^2 \le CNh^{-2}(\frac{4}{(\Delta t)^2}N^{-1} + h^{-2}N^{-1})$ $\Delta t^4 \varepsilon^2 b_n^T K^{-1} b_n$, which implies

$$|s_n|_K \le Ch^{-1}(1 + \frac{\Delta t}{h})\varepsilon\big(\Delta t^{-1}\sqrt{E_n^{\rm ap}} + \Delta t\|g_{n+\frac{1}{2}}\|_{L^2(\Omega)}\big)\,. \tag{7.13}$$

where $E_n^{\rm ap} := \frac{1}{2} p_n^{{\rm ap},T} M^{-1} p_n^{\rm ap} + \frac{1}{2} q_n^{{\rm ap},T} K q_n^{\rm ap}$. Recall that $E_n := \frac{1}{2} p_n^T M^{-1} p_n + \frac{1}{2} q_n^T K q_n$, using $\sqrt{E_n^{\rm ap}} \le \sqrt{E_n} + \sqrt{E_n^{\rm err}}$, we deduce from (7.7) that

$$\sqrt{E_{n+1}^{\rm err}} - \sqrt{E_n^{\rm err}} \le (c_n + z\sqrt{E_n^{\rm err}}) \tag{7.14}$$

with $c_n = C(\frac{1}{h} + \frac{\Delta t}{h^2})\varepsilon\big(\Delta t^{-1}\sqrt{E_n} + \Delta t\|g_{n+\frac{1}{2}}\|_{L^2(\Omega)}\big)$ and $z = C\frac{1}{h}(\frac{1}{\Delta t} + \frac{1}{h})\varepsilon$. Therefore (using $E_0^{\rm err} = 0$) we obtain that $\sqrt{E_n^{\rm err}} \le \sum_{k=0}^{n-1} c_k(1+z)^{n-k-1}$. For $\varepsilon \le \frac{1}{4}C^{-1}\frac{\Delta t}{T}h(\Delta t + h)$, we have $z \le \Delta t/T$, therefore $(1+z)^n \le e^1$ and $\sqrt{E_n^{\rm err}} \le C\sum_{k=0}^{n-1} c_k$. Using $n\Delta t \le T$ and $\|g_{k+\frac{1}{2}}\|_{L^2(\Omega)} \le \|g\|_{L^\infty(0,T,L^2(\Omega))}$, Lemma 7.1 implies that

$$\sqrt{E_n} \le \sqrt{E_0} + \Delta t 2^{-1/2}\sum_{k=0}^{n-1}\|g_{k+\frac{1}{2}}\|_{L^2(\Omega)} \le \sqrt{E_0} + T2^{-1/2}\|g\|_{L^\infty(0,T,L^2(\Omega))}.$$

Using $\sum_{k=0}^{n-1} c_k \le C(\frac{1}{h} + \frac{\Delta t}{h^2})\varepsilon\big(\Delta t^{-1}\sum_{k=0}^{n-1}\sqrt{E_k} + \Delta t\sum_{k=0}^{n-1}\|g_{k+\frac{1}{2}}\|_{L^2(\Omega)}\big)$ we obtain that $\sum_{k=0}^{n-1} c_k \le C\Delta t^{-2}(\frac{1}{h} + \frac{\Delta t}{h^2})\varepsilon T\big(\sqrt{E_0} + T\|g\|_{L^\infty(0,T,L^2(\Omega))}\big)$ and

$$\sqrt{E_n^{\rm err}} \le \Delta t^2\big(\sqrt{E_0} + T\|g\|_{L^\infty(0,T,L^2(\Omega))}\big) \tag{7.15}$$

for $\varepsilon \le C^{-1}\frac{1}{T}\Delta t^4\frac{h}{\Delta t}(\Delta t + h)$. We conclude the proof by observing that $2E_n^{\rm err} = \int_\Omega(\nabla u_n - \nabla u_n^{\rm ap})^T a(\nabla u_n - \nabla u_n^{\rm ap}) + \int_\Omega(v_n - v_n^{\rm ap})^2\mu$.

7.2 Stability

Definition 7.2. *A function $f(t,x)$ is dissipative if $(f(t,y) - f(t,z)),(y - z)) \le 0$ for all y and z. An ODE is contractive if $\|y(t) - z(t)\| \le \|y(s) - z(s)\|$ for every pair of solutions y and z when $t \ge s$. Every ODE with a dissipative right-hand side f is contractive.*

It is easy to see that equation (4.2) is contractive.

Definition 7.3. *A numerical method is B-stable (or contractive) if every pair of numerical solutions u and v satisfy $\|u_{n+1} - v_{n+1}\| \le \|u_n - v_n\|$ for all $n \ge 0$, when solving an IVP with a dissipative f.*

Definition 7.4. *A Runge-Kutta method is algebraically stable if the matrices*

$$B = \mathrm{diag}(b_1,\dots,b_s), \qquad M = BA + A^TB^T - bb^T$$

are nonnegative semidefinite. An algebraically stable Runge Kutta method is B-stable.

7.3 Proof of Theorem 4.1

The implicit Euler scheme (4.3) can be written as

$$(\frac{M}{\Delta t} + K)q_{n+1} = \frac{M}{\Delta t}q_n + f_{n+1}.$$

Using localized gamblets, (4.3) is solved up to error e_n, i.e.

$$(\frac{M}{\Delta t}+K)q_{n+1}^{\rm ap}=\frac{M}{\Delta t}q_n^{\rm ap}+f_{n+1}+(\frac{M}{\Delta t}+K)e_n$$

and e_n satisfies, $e_n^T(\frac{M}{\Delta t}+K)e_n\leq C\varepsilon^2 b_n^T K^{-1}b_n$. Write $\|q\|_\zeta^2:=q^T(\frac{M}{\Delta t}+K)q$.

Lemma 7.5. *It holds true that* $\|q_{n+1}\|_\zeta\leq\|q_n\|_\zeta+\|f_{n+1}\|_{H^{-1}(\Omega)}$.

Proof. Multiplying (4.3) by q_{n+1} and using Young's inequality we obtain that $q_{n+1}^T(\frac{M}{\Delta t}+K)q_{n+1}=q_{n+1}^T\frac{M}{\Delta t}q_n+q_{n+1}^Tf_{n+1}$ and $\|q_{n+1}\|_\zeta^2\leq\frac{1}{2}q_{n+1}^T\frac{M}{\Delta t}q_{n+1}+\frac{1}{2}q_n^T\frac{M}{\Delta t}q_n+\frac{1}{2}q_{n+1}^TKq_{n+1}+\frac{1}{2}f_{n+1}^TK^{-1}f_{n+1}$. Therefore,

$$q_{n+1}^T(\frac{M}{\Delta t}+K)q_{n+1}\leq q_n^T(\frac{M}{\Delta t}+K)q_n+f_{n+1}^TK^{-1}f_{n+1}.$$

which concludes the proof of Lemma 7.5. □

Let $\varepsilon_n:=q_n-q_n^{\rm ap}$, then we have $(\frac{M}{\Delta t}+K)\varepsilon_{n+1}=\frac{M}{\Delta t}\varepsilon_n+(\frac{M}{\Delta t}+K)e_n$ and $(\frac{M}{\Delta t}+K)(\varepsilon_{n+1}-e_n)=\frac{M}{\Delta t}\varepsilon_n$. Multiplying by $\varepsilon_{n+1}-e_n$ and using Young's inequality, we obtain that $(\varepsilon_{n+1}-e_n)^T(\frac{M}{\Delta t}+K)(\varepsilon_{n+1}-e_n)=(\varepsilon_{n+1}-e_n)^T(\frac{M}{\Delta t})\varepsilon_n$ and $\|\varepsilon_{n+1}-e_n\|_\zeta^2\leq\frac{1}{2}(\varepsilon_{n+1}-e_n)^T(\frac{M}{\Delta t}+K)(\varepsilon_{n+1}-e_n)+\frac{1}{2}\varepsilon_n^T(\frac{M}{\Delta t}+K)\varepsilon_n$. Therefore, $\|\varepsilon_{n+1}-e_n\|_\zeta\leq\|\varepsilon_n\|_\zeta$ and

$$\|\varepsilon_{n+1}\|_\zeta\leq\|\varepsilon_n\|_\zeta+\|e_n\|_\zeta\,. \tag{7.16}$$

Since $\|e_n\|_\zeta^2\leq C\varepsilon^2b_n^TK^{-1}b_n$, and $b_n=\frac{M}{\Delta t}q_n^{\rm ap}+f_{n+1}$, we have $b_n^TK^{-1}b_n=q_n^{\rm ap}\frac{M}{\Delta t}K^{-1}\frac{M}{\Delta t}q_n^{\rm ap}+f_{n+1}K^{-1}f_{n+1}$ and $b_n^TK^{-1}b_n\leq q_n^{\rm ap}Kq_n^{\rm ap}(\frac{\lambda_{\max}(M)}{\lambda_{\min}(K)})^2+f_{n+1}K^{-1}f_{n+1}$. Therefore, $\|b_n\|_{K^{-1}}\leq\|q_n^{\rm ap}\|_\zeta+\|f_{n+1}\|_{K^{-1}(\Omega)}$. Using (7.16) we deduce that $\|e_n\|_\zeta\leq C\varepsilon(\|q_n\|_\zeta+\|\varepsilon_n\|_\zeta+\|f_{n+1}\|_{K^{-1}(\Omega)})\leq C\varepsilon\frac{T}{\Delta t}\|g\|_{L^\infty(0,T,H^{-1}(\Omega))}+C\varepsilon\|\varepsilon_n\|_\zeta$. Hence, $\|\varepsilon_{n+1}\|_\zeta\leq C\frac{T}{\Delta t}\varepsilon\|g\|_{L^\infty(0,T,H^{-1}(\Omega))}+(1+C\varepsilon)\|\varepsilon_n\|_\zeta\leq C(\frac{T}{\Delta t})^2\varepsilon\|g\|_{L^\infty(0,T,H^{-1}(\Omega))}$, which finishes the proof of Theorem 4.1.

Acknowledgements. H. Owhadi gratefully acknowledges the support of the Air Force Office of Scientific Research and the DARPA EQUiPS Program under award number FA9550-16-1-0054 (Computational Information Games). L. Zhang gratefully acknowledges the support of the National Natural Science Foundation of China grant 11471214 and the One Thousand Plan of China for young scientists. The authors also thank two anonymous referees for comments and suggestions.

References

[1] A. Averbuch, G. Beylkin, R. Coifman, P. Fischer, and M. Israeli. Adaptive solution of multidimensional PDEs via tensor product wavelet decomposition. Int. J. Pure Appl. Math., 44(1):75–115, 2008.

[2] I. Babuška and R. Lipton. Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. Multiscale Model. Simul., 9:373–406, 2011.

[3] I. Babuška and J. E. Osborn. Generalized finite element methods: their performance and their relation to mixed methods. SIAM J. Numer. Anal., 20(3):510–536, 1983.

[4] E. Bacry, S. Mallat, and G. Papanicolaou. A wavelet space-time adaptive scheme for partial differential equations. In Progress in wavelet analysis and applications (Toulouse, 1992), pages 677–682. Frontières, Gif-sur-Yvette, 1993.

[5] R. E. Bank, W. M. Coughran, W Fichtner, E. H. Grosse, D. J. Rose, and R. K. Smith. Transient simulation of silicon devices and circuits. IEEE Transactions on Computer-Aided Design, 32(10):1992–2007, 1985.

[6] R. E. Bank, T. F. Dupont, and H. Yserentant. The hierarchical basis multigrid method. Numer. Math., 52(4):427–458, 1988.

[7] M. Bebendorf. Hierarchical matrices, volume 63 of Lecture Notes in Computational Science and Engineering. Springer-Verlag, Berlin, 2008. A means to efficiently solve elliptic boundary value problems.

[8] G. Ben Arous and H. Owhadi. Multiscale homogenization with bounded ratios and anomalous slow diffusion. Comm. Pure Appl. Math., 56(1):80–113, 2003.

[9] L. Berlyand and H. Owhadi. Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast. Archives for Rational Mechanics and Analysis, 198(2):677–721, 2010.

[10] G. Beylkin and N. Coult. A multiresolution strategy for reduction of elliptic PDEs and eigenvalue problems. Appl. Comput. Harmon. Anal., 5(2):129–155, 1998.

[11] P.D. Boom and Zingg D.W. High-order implicit time-marching methods based on generalized summation-by-parts operators. SIAM Journal on Scientific Computing, 37(6):A2682–A2709, 2015.

[12] A. Brandt. Multi-level adaptive technique (MLAT) for fast numerical solutions to boundary value problems. In Proc. 3rd Int'l Conf. Numerical Methods in Fluid Mechanics, 1973. Lecture Notes in Physics 18.

[13] L. V. Branets, S. S. Ghai, L. L., and X.-H. Wu. Challenges and technologies in reservoir modeling. Commun. Comput. Phys., 6(1):1–23, 2009.

[14] M. E. Brewster and G. Beylkin. A multiresolution strategy for numerical homogenization. Appl. Comput. Harmon. Anal., 2(4):327–349, 1995.

[15] F.-X. Briol, C. J. Oates, M. Girolami, M. A. Osborne, and D. Sejdinovic. Probabilistic integration: A role for statisticians in numerical analysis? arXiv:1512.00933, 2015.

[16] J. C. Butcher. Numerical Methods for Ordinary Differential Equations. John Wiley & Sons, Ltd, 2008.

[17] J. M. Carnicer, W. Dahmen, and J. M. Peña. Local decomposition of refinable spaces and wavelets. Applied and Computational Harmonic Analysis, 3(2):127–153, 1996.

[18] G. Chiavassa and J. Liandrat. A fully adaptive wavelet algorithm for parabolic partial differential equations. Appl. Numer. Math., 36(2-3):333–358, 2001.

[19] O. A. Chkrebtii, D. A. Campbell, B. Calderhead, and M. A. Girolami. Bayesian solution uncertainty quantification for differential equations. Bayesian Analysis. arXiv:1306.2365, 2016.

[20] J. Cockayne, C. J. Oates, T. Sullivan, and M. A. Girolami. Probabilistic meshless methods for partial differential equations and bayesian inverse problems. arXiv:1605.07811, 2016.

[21] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. Communications on Pure and Applied Mathematics, 45(5):485–560, 1992.

[22] Albert Cohen. Adaptive methods for pde's: wavelets or mesh refinement? arXiv preprint math/0212414, 2002.

[23] P. R. Conrad, M. Girolami, S. Särkä, A. Stuart, and K. Zygalakis. Probability measures for numerical solutions of differential equations. Statistics and Computing, arXiv:1512.00933, pages 1–18, 2016.

[24] W. Dahmen and A. Kunoth. Adaptive wavelet methods for linear-quadratic elliptic control problems: convergence rates. SIAM J. Control Optim., 43(5):1640–1675, 2005.

[25] M. Desbrun, R. Donaldson, and H. Owhadi. Modeling across scales: Discrete geometric structures in homogenization and inverse homogenization. Reviews of Nonlinear Dynamics and Complexity. Special issue on Multiscale Analysis and Nonlinear Dynamics., 2012.

[26] P. Diaconis. Bayesian numerical analysis. In Statistical Decision Theory and Related Topics, IV, Vol. 1 (West Lafayette, Ind., 1986), pages 163–175. Springer, New York, 1988.

[27] Jean Dieudonné et al. On biorthogonal systems. The Michigan Mathematical Journal, 2(1):7–20, 1953.

[28] M. Dorobantu and B. Engquist. Wavelet-based numerical homogenization. SIAM J. Numer. Anal., 35(2):540–559 (electronic), 1998.

[29] Hosea M. E. and Shampine L. F. Analysis and implementation of tr-bdf2. Applied Numerical Mathematics, 20:21–37, 1996.

[30] W. E and B. Engquist. The heterogeneous multiscale methods. Commun. Math. Sci., 1(1):87–132, 2003.

[31] Y. Efendiev, V. Ginting, T. Hou, and R. Ewing. Accurate multiscale finite element methods for two-phase flow simulations. J. Comput. Phys., 220(1):155–174, 2006.

[32] Bjorn Engquist, Stanley Osher, and Sifen Zhong. Fast wavelet based algorithms for linear evolution equations. SIAM Journal on Scientific Computing, 15(4):755–775, 1994.

[33] Y. A. Erlangga, C. W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous helmholtz problems. SIAM J. Sci. Comput., 27(4):1471–1492, 2006.

[34] R. P. Fedorenko. A relaxation method of solution of elliptic difference equations. Ž. Vyčisl. Mat. i Mat. Fiz., 1:922–927, 1961.

[35] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. J. Comput. Phys., 73(2):325–348, 1987.

[36] W. Hackbusch. A fast iterative method for solving Poisson's equation in a general region. In Numerical treatment of differential equations (Proc. Conf., Math. Forschungsinst., Oberwolfach, 1976), pages 51–62. Lecture Notes in Math., Vol. 631. Springer, Berlin, 1978.

[37] W. Hackbusch, L. Grasedyck, and S. Börm. An introduction to hierarchical matrices. In Proceedings of EQUADIFF, 10 (Prague, 2001), volume 127, pages 229–241, 2002.

[38] E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration, volume 31 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[39] E. Hairer and G. Wanner. Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems. Springer Series in Computational Mathematics. Springer-Verlag, Berlin, second revised edition edition, 1996.

[40] P. Hennig. Probabilistic interpretation of linear solvers. SIAM Journal on Optimization, 25(1):234–260, 2015.

[41] P. Hennig, M. A. Osborne, and M. Girolami. Probabilistic numerics and uncertainty in computations. Proc. A., 471(2179):20150142, 17, 2015.

[42] K. Ho and L. Ying. Hierarchical interpolative factorization for elliptic operators: Differential equations. Communiations on Pure and Applied Mathematics, 69(8):1415–1451, 2016.

[43] M. Hochbruck and T. Pazur. Implicit runge-kutta methods and discontinuous galerkin discretizations for linear maxwell's equations. SIAM J. Numer. Anal., 53(1):485–507, 2015.

[44] M. Hochbruck, T. Pazur, A. Schulz, E. Thawinan, and Wieners C. Efficient time integration for discontinuous galerkin approximations of linear wave equations. ZAMM, 95(3):237–259, 2015.

[45] T. H. Hou and P. Liu. Optimal local multi-scale basis functions for linear elliptic equations with rough coefficient. Discrete and Continuous Dynamical Systems, 36(8):4451–4476, 2016.

[46] T. Y. Hou and X. H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. J. Comput. Phys., 134(1):169–189, 1997.

[47] G. S. Kimeldorf and G. Wahba. A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. Ann. Math. Statist., 41:495–502, 1970.

[48] F. M. Larkin. Gaussian measure in Hilbert space and applications in numerical analysis. Rocky Mountain J. Math., 2(3):379–421, 1972.

[49] Peter D Lax. Integrals of nonlinear equations of evolution and solitary waves. Communications on pure and applied mathematics, 21(5):467–490, 1968.

[50] C. Lubich and A. Ostermann. Multigrid dynamic iteration for parabolic equations. BIT, 27(2):216–234, 1987.

[51] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. Mathematics of Computation, 83(290):2583–2603, 2014.

[52] J. Mandel, M. Brezina, and P. Vaněk. Energy optimization of algebraic multigrid bases. Computing, 62(3):205–228, 1999.

[53] Nicola Marzari, Arash A. Mostofi, Jonathan R. Yates, Ivo Souza, and David Vanderbilt. Maximally localized wannier functions: Theory and applications. Rev. Mod. Phys., 84:1419–1475, Oct 2012.

[54] C. A. Micchelli and T. J. Rivlin. A survey of optimal recovery. In Optimal Estimation in Approximation Theory, pages 1–54. Springer, 1977.

[55] A. S. Nemirovsky. Information-based complexity of linear operator equations. J. Complexity, 8(2):153–175, 1992.

[56] E. Novak and H. Woźniakowski. Tractability of multivariate problems. Volume II: Standard information for functionals, volume 12 of EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2010.

[57] A. O'Hagan. Bayes-Hermite quadrature. J. Statist. Plann. Inference, 29(3):245–260, 1991.

[58] A. O'Hagan. Some Bayesian numerical analysis. In Bayesian Statistics, 4 (Peñíscola, 1991), pages 345–363. Oxford Univ. Press, New York, 1992.

[59] H. Owhadi. Anomalous slow diffusion from perpetual homogenization. Ann. Probab., 31(4):1935–1969, 2003.

[60] H. Owhadi. Averaging versus chaos in turbulent transport? Comm. Math. Phys., 247(3):553–599, 2004.

[61] H. Owhadi. Bayesian numerical homogenization. Multiscale Model. Simul., 13(3):812–828, 2015.

[62] H. Owhadi. Multigrid with Rough Coefficients and Multiresolution Operator Decomposition from Hierarchical Information Games. SIAM Rev., 59(1):99–149, 2017. arXiv:1503.03467, 2015.

[63] H. Owhadi and C. Scovel. Brittleness of Bayesian inference and new Selberg formulas. Communications in Mathematical Sciences, 14:83–145, 2016. arXiv:1304.7046.

[64] H. Owhadi and C. Scovel. Qualitative robustness in bayesian inference. arXiv:1411.3984, 2016.

[65] H. Owhadi and C. Scovel. Towards Machine Wald. In R. Ghanem, D. Higdon, and H. Owhadi, editors, Handbook for Uncertainty Quantification, pages 1–35. Springer International Publishing, 2017. arXiv:1508.02449.

[66] H. Owhadi and C. Scovel. Universal scalable robust solvers from computational information games and fast eigenspace adapted multiresolution analysis. arXiv:1703.10761, 2017.

[67] H. Owhadi, C. Scovel, and T. J. Sullivan. Brittleness of Bayesian Inference under finite information in a continuous world. Electronic Journal of Statistics, 9:1–79, 2015. arXiv:1304.6772.

[68] H. Owhadi, C. Scovel, and T. J. Sullivan. On the Brittleness of Bayesian Inference. SIAM Review, 57(4):566—-582, 2015.

[69] H. Owhadi and L. Zhang. Homogenization of parabolic equations with a continuum of space and time scales. SIAM Journal on Numerical Analysis, 46(1):1–36, 2007.

[70] H. Owhadi and L. Zhang. Metric-based upscaling. Comm. Pure Appl. Math., 60(5):675–723, 2007.

[71] H. Owhadi and L. Zhang. Homogenization of the acoustic wave equation with a continuum of scales. Computer Methods in Applied Mechanics and Engineering, 198(3–4):397–406, 2008.

[72] H. Owhadi and L. Zhang. Localized bases for finite dimensional homogenization approximations with non-separated scales and high-contrast. SIAM Multiscale Modeling & Simulation, 9:1373–1398, 2011. arXiv:1011.0986.

[73] H. Owhadi, L. Zhang, and L. Berlyand. Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization. ESAIM Math. Model. Numer. Anal., 48(2):517–552, 2014.

[74] E. W. Packel. The algorithm designer versus nature: a game-theoretic approach to information-based complexity. J. Complexity, 3(3):244–257, 1987.

[75] Paris Perdikaris, Daniele Venturi, and George Em Karniadakis. Multifidelity information fusion algorithms for high-dimensional systems and massive data sets. SIAM Journal on Scientific Computing, 38(4):B521–B538, 2016.

[76] D. Peterseim. Variational multiscale stabilization and the exponential decay of fine-scale correctors. arXiv:1505.07611, 2015.

[77] H. Poincaré. Calcul des probabilités. Georges Carrés, Paris, 1896.

[78] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Inferring solutions of differential equations using noisy multi-fidelity data. Journal of Computational Physics, 335:736–746, 2017.

[79] K. Ritter. Average-case analysis of numerical problems, volume 1733 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2000.

[80] J. W. Ruge and K. Stüben. Algebraic multigrid. In Multigrid methods, volume 3 of Frontiers Appl. Math., pages 73–130. SIAM, Philadelphia, PA, 1987.

[81] A. Sard. Linear approximation. American Mathematical Society, Providence, R.I., 1963.

[82] F. Schäfer, T. J. Sullivan, and H. Owhadi. Compression, inversion, and approximate PCA of dense kernel matrices at near-linear computational complexity. To appear, 2017.

[83] Florian Schäfer, TJ Sullivan, and Houman Owhadi. Compression, inversion, and approximate PCA of dense kernel matrices at near-linear computational complexity. arXiv preprint arXiv:1706.02205, 2017.

[84] M. Schober, D. K. Duvenaud, and P. Hennig. Probabilistic ODE solvers with Runge-Kutta means. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger, editors, Advances in Neural Information Processing Systems 27, pages 739–747. Curran Associates, Inc., 2014.

[85] J. E. H. Shaw. A quasirandom approach to integration in Bayesian statistics. Ann. Statist., 16(2):895–914, 1988.

[86] J. Skilling. Maximum Entropy and Bayesian Methods: Seattle, 1991, chapter Bayesian Solution of Ordinary Differential Equations. Springer Netherlands, Dordrecht, 1992.

[87] R. Stevenson. Adaptive wavelet methods for solving operator equations: an overview. In Multiscale, Nonlinear and Adaptive Approximation, pages 543–597. Springer, 2009.

[88] R. Sudarshan. Operator-adapted finite element wavelets: Theory and applications to a posteriori error estimation and adaptive computational modeling. ProQuest LLC, Ann Arbor, MI, 2005. Thesis (Ph.D.)–Massachusetts Institute of Technology.

[89] A. V. Sul′din. Wiener measure and its applications to approximation methods. I. Izv. Vyssˇ. Učebn. Zaved. Matematika, 1959(6 (13)):145–158, 1959.

[90] W. Sweldens. The lifting scheme: A construction of second generation wavelets. SIAM journal on mathematical analysis, 29(2):511–546, 1998.

[91] W. Symes. Transfer of approximation and numerical homogenization of hyperbolic boundary value problems with a continuum of scales. TR12-20 Rice Tech Report, 2012.

[92] M. Tao, H. Owhadi, and J. E. Marsden. Nonintrusive and structure preserving multiscale integration of stiff ODEs, SDEs, and Hamiltonian systems with hidden slow dynamics via flow averaging. Multiscale Model. Simul., 8(4):1269–1324, 2010.

[93] M. Tao, H. Owhadi, and J. E. Marsden. From efficient symplectic exponentiation of matrices to symplectic integration of high-dimensional Hamiltonian systems with slowly varying quadratic stiff potentials. Appl. Math. Res. Express. AMRX, 2:242–280, 2011.

[94] M. Tao, H. Owhadi, and Jerrold E. Marsden. Space-time FLAVORS: finite difference, multisymlectic, and pseudospectral integrators for multiscale PDEs. Dyn. Partial Differ. Equ., 8(1):21–45, 2011.

[95] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski. Information-based complexity. Computer Science and Scientific Computing. Academic Press, Inc., Boston, MA, 1988. With contributions by A. G. Werschulz and T. Boult.

[96] Lent J. V. and S. Vandewalle. Multigrid methods for implicit runge-kutta and boundary value method discretizations of parabolic pdes. SIAM J. Sci. Comput., 27(1):67–92, 2005.

[97] S. Vandewalle. Parallel Multigrid Waveform Relaxation for Parabolic Problems. Vieweg+Teubner Verlag, Stuttgart, 1993.

[98] P. S. Vassilevski. Multilevel preconditioning matrices and multigrid V-cycle methods. In Robust multi-grid methods (Kiel, 1988), volume 23 of Notes Numer. Fluid Mech., pages 200–208. Vieweg, Braunschweig, 1989.

[99] P. S. Vassilevski. On two ways of stabilizing the hierarchical basis multilevel methods. SIAM Rev., 39(1):18–53, 1997.

[100] P. S. Vassilevski and J. Wang. Stabilizing the hierarchical basis by approximate wavelets. I. Theory. Numer. Linear Algebra Appl., 4(2):103–126, 1997.

[101] A. Wald. Statistical decision functions which minimize the maximum risk. Ann. of Math. (2), 46:265–280, 1945.

[102] W. L. Wan, Tony F. Chan, and Barry Smith. An energy-minimizing interpolation for robust multigrid methods. SIAM J. Sci. Comput., 21(4):1632–1649, 1999/00.

[103] X. Wang. Transfer-of-approximation approaches for subgrid modeling. PhD thesis, Rice University, 2012.

[104] Gregory H. Wannier. The structure of electronic excitation levels in insulating crystals. Phys. Rev., 52:191–197, Aug 1937.

[105] C. D. White and R. N. Horne. Computing absolute transmissibility in the presence of finescale heterogeneity. SPE Symposium on Reservoir Simulation, page 16011, 1987.

[106] C. Wieners. A geometric data structure for parallel finite elements and the application to multigrid methods with block smoothing. Computing and Visualization in Science, 13:161–175, 2010.

[107] H. Woźniakowski. Probabilistic setting of information-based complexity. J. Complexity, 2(3):255–269, 1986.

[108] H. Woźniakowski. What is information-based complexity? In Essays on the complexity of continuous problems, pages 89–95. Eur. Math. Soc., Zürich, 2009.

[109] J. Xu and L. Zikatanov. On an energy minimizing basis for algebraic multigrid methods. Comput. Vis. Sci., 7(3-4):121–127, 2004.

[110] I. Yavneh. Why multigrid methods are so efficient. Computing in Science and Engg., 8(6):12–22, November 2006.

[111] L. Ying, G. Biros, and D. Zorin. A kernel-independent adaptive fast multipole method in two and three dimensions. Journal of Computational Physics, 196(2):591–626, 2004.

[112] H. Yserentant. On the multilevel splitting of finite element spaces. Numer. Math., 49(4):379–412, 1986.